>> DAVID BAHAR: THANK YOU EVERYONE, AND I'M VERY MUCH LOOKING FORWARD TO OUR DISCUSSION TODAY. I WANTED TO CHECK WITH THE ROOM HERE IN TERMS OF EVERYONE'S COMFORT AND FAMILIARITY WITH THE TERM ASR, WHO KNOWS WHAT ASR WORKS? BY SHOW OF HANDS. WHO KNOWS IN-DEPTH HOW THE TECHNOLOGY REALLY FUNCTIONS BY SHOW OF HANDS. I DON'T SEE A HAND. OKAY. WELL, THAT'S A GOOD TRANSITION POINT. I SEE SOMEBODY OVER THERE. WELL, JEFF, WELL, YEAH, YOU WOULD KNOW, WOULDN'T YOU. (LAUGHTER).

THAT'S A GOOD TRANSITION FOR US, THEN, FOR THE FIRST QUESTION. MAYBE WE'LL START WITH VISAR. COULD YOU GIVE US AN OVERVIEW OF JUST EXACTLY WHAT ASR IS AND WHAT I WAS USING IN MICROSOFT TO TRANSLATE A PRESENTATION.

>> VISAR BERISHA: SURE. I'D BE HAPPY TO DO THAT. I ACTUALLY HAVE SOME SLIDES THAT I CAN GO THROUGH. WELL -- THANK YOU. I'LL GO AHEAD AND DRIVE. I THINK MANY OF US ARE FAMILIAR WITH THE VARIOUS TECHNOLOGIES THAT ARE OUT THERE, THE VOICE TECHNOLOGIES. THESE HAVE BECOME A PART OF SEVERAL HOUSEHOLDS, MANY HOUSEHOLDS, SO YOU HAVE CORTANA, ALEXA, SIRI, GOOGLE, AND ALL OF THIS WHAT UNDER LIES IS THE SPEECH RECOGNITION. SO WHEN SOMEONE SPEAKS AN ACOUSTIC SIGNAL IS GENERATED. SO THE DEVICE IS ABLE TO RECORD, BECAUSE THERE'S A MICROPHONE ON THE DEVICE, THE SIGNAL, AND TO CONVERT IT TO A STRING OF WORDS.

SO AUTOMATIC SPEECH RECOGNITION IS THE PROCESS BEHIND WHICH THAT OCCURS, TAKING SPEECH AND CONVERTING IT TO TEXT. AND SO WHEN WE'RE EVALUATING AUTOMATIC SPEECH RECOGNITION, WHAT WE'LL REALLY TRYING TO DO IS IDENTIFY AN ALGORITHM, IDENTIFY A PROCESS BY WHICH WE AUTOMATICALLY TRANSCRIBE INCOMING TEXT MAKING AS ~~FIE~~ FEW TRANSCRIPTION ERRORS AS POSSIBLE AND DOING THIS IN A REASONABLE TIME FRAME. IF IT TOOK YOUR ALEXA10 SECONDS TO PROVIDE A RESPONSE, THAT WOULD SIGNIFICANTLY HINDER THE PERCEIVED SUCCESS OF THE CONVERSATION, SO THE LATENCY BECOMES A VERY IMPORTANT PARAMETER HERE.

AND FINALLY IT CERTAINLY WOULD BE VERY NICE TO BE ABLE TO DO THIS IN A WAY THAT ~~PROVIDES~~ PRESERVES THE PRIVACY OF THE USER USING THIS TECHNOLOGY. SO THESE ARE SORT OF OUR DESIRABLES FOR A SPEECH RECOGNITION SYSTEM. AND WHAT I'LL TALK ABOUT HERE ~~SUSPECT~~ IS AN OVERVIEW OF HOW THE VARIOUS COMPANIES THAT HAVE DEVELOPED THESE DEVICES DEVELOPED THESE ALGORITHMS.

OKAY. SO ONE THING THAT WE CAN PERHAPS DO IS TAKE INSPIRATION FROM HOW HUMANS DO THIS. AND SO THERE WILL BE A TEST AFTERWARDS. (LAUGHTER) SO, PLEASE, YOU KNOW, PAY CAREFUL ATTENTION AND TAKE NOTES. SO HOW DO HUMANS DO THIS?

WELL, AS SOON AS WE GENERATE A SPEECH SIGNAL BY SPEAKING, THIS REALLY RESULTS IN SORT OF DISTURBANCE OF THE AIR IN FRONT OF US, SO SPEECH IS NOTHING MORE THAN A PRESSURE WAVE, AND THIS PRESSURE WAVE, THIS THING THAT CAN RECORD WITH THE MICROPHONE ALSO HITS THE EARDRUM OF THE LISTENER, AND THE EARDRUM AND THE COCHLEA DOES SOMETHING VERY INTERESTING. IT TAKES THIS PRESSURE WAVE AND CONVERTS IT TO A REPRESENTATION THAT ALLOWS DIFFERENT TYPES OF SOUNDS THAT MAKE UP SPEECH TO STAND OUT. SO INTERNALLY, AS THE EARDRUM AND ~~THE FOLLOW ON THE~~ COCHLEA PROCESS THE SIGNAL, IT DOES SOMETHING TO IT SO THAT THE S'S STAND OUT FROM THE A'S, FROM THE E'S, FROM THE U'S, THAT MAKE UP THIS SIGNAL, ~~THESE~~ THE VERY SOUNDS WE CALL PHONEMES. ~~SO YOU PUT WORDS TOGETHER TO GENERATE SENTENCES AND SO ON.~~

FOLLOWING THIS PROCESSING BY THE COCHLEA, THIS NEW REPRESENTATION THIS DECOMPOSITION OF THIS PRESSURE WAVE INTO A SET OF PARAMETERS THAT HELP THE BRAIN DISTINGUISH DIFFERENT SOUNDS FROM EACH OTHER IS SENT ON TO THE BRAIN. AND VARIOUS NEURONAL NETWORKS ARE ABLE TO IDENTIFY THE BEGINNINGS AND ENDS OF SOUNDS AND THE BEGINNING OF WORDS AND SENTENCES AND SO ON. AND THIS IS HOW WE COMMUNICATE THROUGH SPEECH.

THIS ENDS UP BEING A VERY ROBUST SYSTEM, AND SOMETHING THAT IF WE CAN MIMIC ALGORITHMICLY WOULD BE VERY BENEFICIAL. SO HUMANS ARE ABLE TO ADAPT THEIR LISTENING STRATEGIES AND THE WAY THEY MAKE SENSE OF SOUNDS ON THE FLY. IF THERE'S A NOISY ENVIRONMENT, THE WAY THAT YOU LISTEN IS DIFFERENT THAN IF IT'S AN ENVIRONMENT WITH NO NOISE.

IF YOU KNOW THE PERSON, AS YOU CONTINUALLY INTERACT WITH THEM, YOU MODIFY THE WAY THAT YOU LISTEN TO THAT PERSON, AND SO OVER TIME YOU END UP RECOGNIZING MORE OF THE WORDS THAT THEY'RE SAYING. THIS IS WHY, FOR EXAMPLE, YOU KNOW, WE CAN UNDERSTAND OUR CHILDREN, BUT NO ONE ELSE CAN WHEN THEY'RE VERY YOUNG.

AND IT'S REALLY MORE BROADLY THAN THAT, ROBUST TO DIFFERENT LEVELS OF BACKGROUND NOISE, DIFFERENT ~~LEFLZ~~ LEVELS OF REVERBERATION, SO IF YOU HAVE SMALLER ROOMS OR LARGER

ROOMS, THAT RESULTS IN DIFFERENT REVERBERATION PATTERNS, DIALECTS, AND SPEECH PATTERNS ~~DEVELOPING~~ FROM INDIVIDUALS WITH PARKINSONS DISEASE.

SO IT WOULD BE NICE IF WE COULD GENERATE COMPUTER ALGORITHMS THAT WERE EQUALLY ROBUST. WHEN SPEECH TECHNOLOGY BEGAN, WHEN SPEECH RECOGNITION STARTED AS A FIELD, THERE WAS SOME ATTEMPT TO DO THIS BY REALLY CODIFYING SOME RULES THAT WE KNOW HUMANS USE IN ORDER TO PERCEIVE SPEECH. AND BROADLY SPEAKING, THIS SORT OF WORKS IN THREE DIFFERENT PROCESSES. IN THE FIRST, THERE IS SOME ACOUSTIC WAVE FORM THAT'S GENERATED, AN ACOUSTIC WAVEFORM IS BASICALLY WHAT'S RECORDED IN A MICRO FROM PHONE FROM A PERSON SPEAKING. THAT'S TRANSFORMED INTO A REPRESENTATION, MUCH LIKE THE COCHLEA DOES, TRANSFORMS THE SPEECH SIGNAL INTO REPRESENTATION THAT ALLOWS DIFFERENT SPEECH SOUNDS TO STAND OUT FROM EACH OTHER, SO THAT'S THE MIDDLE PANEL THERE. AND THEN THIS SEQUENCE IS CONVERTED INTO SOUNDS, WORDS AND SENTENCES AUTOMATICALLY BY THE COMPUTER.

NOW, THIS SOUNDS EASY ENOUGH, AND, YOU KNOW, IT'S CERTAINLY EASY TO EXPLAIN AT THIS LEVEL OF ABSTRACTION, BUT IT'S ACTUALLY QUITE CHALLENGING. BUT INHERENT IN THIS PROCESS, YOU NEED SEVERAL SUB PROCESSES, SUB ALGORITHMS THAT MAKE THIS ROBUST CONDITION. SO IT WOULD BE NICE IF PEOPLE ONLY SPOKE IN ENVIRONMENTS WHERE THERE'S NO BACKGROUND NOISE, BUT THAT'S NOT THE CASE, SO WE HAVE TO DEVELOP ALGORITHMS IN SPEECH TO SEPARATE THE NOISE FROM THE SPEECH SIGNAL THAT YOU'RE INTERESTED IN, AND THIS IS, OF COURSE, VERY CHALLENGING WHEN THERE'S OTHER BACKGROUND SPEAKERS; IN RESTAURANTS, FOR EXAMPLE.

THERE'S LOTS OF ~~VAEBLT~~ VARIABILITY IN HUMAN SPEECH, THE WAY THAT I SPEAK IS DIFFERENT HAVE THE WAY THAT OTHER PEOPLE SPEAK, AND SO THESE ALGORITHMS HAVE TO BE ROBUST, AND TO BE ABLE TO UNDERSTAND FOLKS WITH DIFFERENT DIALECTS, AND DIFFERENT PRECISION, SO THESE ALGORITHMS HAVE TO BE ROBUST IN THE DIFFERENT SYSTEMS. THERE'S SOME WORDS THAT SOUND DIFFERENT THAN THE OTHER, TO ALLUSION, ELUSION, AND ILLUSION, UNLESS YOU KNOW THE CONTEXT, IT BECOMES DIFFICULT TO DECIPHER THESE.

AND THEN ~~SHG~~ OF COURSE IN CONVERSATIONAL SETTINGS, PEOPLE SPEAK WITH CONVERSATIONAL FILLERS, AND SO IF I'M TRYING THIS THINK OF THE NEXT WORD TO SAY NEXT, I MIGHT SAY "UH," OR IF I BEGIN A SENTENCE, I THINK THAT'S THE WRONG THING TO SAY AND I

GO BACK AND ~~START~~ RESTART, THESE ARE WHAT WE CALL DISFLUENCIES, AND THESE BECOME VERY PROBLEMATIC FOR SPEECH RECOGNITION SYSTEMS.

SO HOW DO THESE ALGORITHMS THAT RESIDE AT GOOGLE AND AT MICROSOFT AND AT AMAZON, HOW DO THEY SOLVE THESE ~~CHAL EDGES?~~ CHALLENGES VERY BROADLY SPEAKING, ~~IT'S~~ AT THE SORT OF THE 10,000-FOOT LEVEL~~.~~, THE ANSWER IS THROUGH DATA. ~~SOF~~SO THEY'VE DEVELOPED DATASETS OF PEOPLE SPEAKING ~~IN~~AND CORRESPONDING TRANSCRIPTIONS ON A MASSIVE SCALE.

WHAT I HAVE HERE IS JUST A TITLE FROM A PAPER THAT WAS PUBLISHED LAST YEAR ON DEVELOPING AN ACOUSTIC MODEL FOR SPEAK RECOGNITION WITH ONE MILLION HOURS OF SPEECH. THAT'S REALLY A SCALE THAT WE HAVEN'T SEEN UNTIL NOW.

SO WHEN DAVID WAS TALKING ABOUT HIS MICROSOFT OFFICE EXPERIENCE, I CAN CERTAINLY ALMOST GUARANTEE THAT THE MODEL THAT UNDERLIES THAT ALGORITHM WAS NOT CALIBRATED WITH A MILLION HOURS OF SPEECH.

SPEECH BY ITSELF IS ONLY -- IS HELPFUL UP TO A POINT. WHAT IS REALLY DONE INTERNALLY TO DEVELOP THESE ALGORITHMS IS YOU HAVE A SPEECH SIGNAL, SO OF YOU HAVE THE RECORDED SPEECH, PERHAPS FROM FOLKS USING ALEXA, THOSE RECORDINGS GOING UP TO THE CLOUD, AND THEN A CORRESPONDING TRANSCRIPTION. SO YOU HAVE THE TRANSCRIPTION FOR THE SAME SPEECH SIGNAL.

IF YOU HAVE DATA LIKE THIS ON A LARGE SCALE, THEN YOU CAN DEVELOP ALGORITHMS TO LEARN THE SORT OF RELATIONSHIP BETWEEN THIS ACOUSTIC, THE RECORDED SPEECH, AND THE TRANSCRIPTS. YOU NEED DATA ON A MASSIVE SCALE, BECAUSE AS I MENTIONED, SPEECH IS VERY VARIABLE, SO YOU HAVE TO HAVE DIFFERENT SPEAKERS, DIFFERENT NOISE CONDITIONS, AND SIMILARLY ON THE TRANSCRIPT SIDE, YOU HAVE TO MAKE SURE THAT YOU CAPTURE THE CONTEXT OF THE RIGHT CONVERSATIONS, AND YOU CAPTURE DISFLUENCIES, AND OTHER SIMILAR, SORT OF, HUMAN ERRORS IN THE TRANSCRIPT.

SO IF YOU HAVE DATA, YOU KNOW, HUNDREDS OF THOUSANDS OF DATA TO SCALE, YOU CAN DEVELOP MACHINE LEARNING MODEL, MACHINE LEARNING IS A BUZZ WORD NOW, BUT NOTHING MORE THAN MATHEMATICAL ALGORITHMS THAT TAKE THE RECORDED SPEECH AS AN INPUT AND PRODUCE TEXT AT THE OUTPUT.

SO AFTER DATA IS COLLECTED AT THIS SCALE, YOU MIGHT IMAGINE THAT THE ALGORITHMS THEMSELVES WILL BE VERY ROBUST, BECAUSE WE DEVELOPED THIS MODEL NOW ON HUNDREDS OF THOUSANDS OF HOURS OF SPEECH AND PERHAPS EVEN A MILLION HOURS OF SPEECH.

SO IF YOU HAVE AN INPUT TO THE MODEL THAT TAKES THE RECORDED SPEECH AND PRODUCES WORDS, LET'S SAY THE INPUT HERE, I DON'T THINK THE AUDIO WORKS, BUT THIS IS ME SAYING "NEW AGE MEDICINE WORKS VERY WELL." IF YOU PROCESS THAT THROUGH THE -- THE SYSTEM THAT TAKES THE RECORDED SPEECH AND PRODUCES WORDS, YOU MAY GET SOMETHING LIKE THIS: "NOON AGE MADISON WORKS VERY WELD ~~S~~" SO IT VERY LIKELY THAT THAT PRECISE SEQUENCE OF WORDS HAS ~~NEFR~~ NEVER BEEN ~~MUTTERED~~UTTERED UNTIL NOW, SO THERE HAS TO BE SOME MECHANISM BY WHICH THESE ALGORITHMS CAN TAKE CONTEXT INTO ACCOUNT.

BECAUSE WE HAVE TRANSCRIPTS, AND BECAUSE WE HAVE DATA ON A VERY LARGE SCALE, WE CAN DESIGN ALGORITHMS THAT GO IN AND SUPERIMPOSE RULES BASED ON THE LANGUAGE. SO IF YOU ANALYZE MILLIONS OF HOURS OF DATA, YOU WILL BE ABLE TO RECOGNIZE THAT NOON AGE MADISON WORKS VERY WELD MAKES NO SENSE, BUT NEW AGE MEDICINE WORKS VERY WELL, IS HE MANTICLY AT LEAST MAKES SENSE. SEW THERE'S THIS SECONDARY STEP WHERE THERE'S A REFINEMENT TO THE ACOUSTIC MODEL BASED ON WHAT MAKES SENSE SEMANTICALLY AND WHAT DOESN'T.

SO SO THESE MODELS NOW THAT THEY'VE BEEN DEVELOPED BY THE LARGE TECHNOLOGY COMPANIES AND THE WAY THAT THEY'RE PROVIDED TO THE PUBLIC AT LARGE, FIRST, IF YOU END UP USING YOUR DEVICES, YOU INTERACT WITH THESE ALGORITHMS ALL THE TIME. BUT IF YOU'RE INTERESTED IN DEVELOPING ~~YUR~~ YOUR OWN TOOLS, THEY OPEN UP THESE ALGORITHMS TO DIFFERENT DEVELOPERS FOR VARIOUS COMMERCIAL APPLICATIONS. SO THIS IS A LOT TRANSCRIBED HERE, ONE OF THE APPS THAT I KNOW IS QUITE POPULAR.

THE WAY THAT THIS WORKS IS ONCE YOU OPEN UP THE APP ON THE PHONE, YOU SPEAK INTO IT, THE AUDIO IS TRANSMITTED TO THE SERVER, AND IS HE SERVERS, THERE'S SOME AUTOMATIC SPEECH RECOGNITION, SO THE AUDIO IS TRANSCRIBED. IN CERTAIN CASES -- I'M NOT SURE IF THIS IS THE CASE WITH WHAT I TRANSCRIBED -- BUT IN SOME CASES SOME OF THE AUDIO IS ACTUALLY STORED ON THEIR SERVERS FOR PROCESSING FOR IMPROVEMENT OF THE ALGORITHMS, ~~SXN~~ THEN FINALLY THE TEXT FROM THE TRANSCRIPTION IS RETURNED BACK TO THE USER.

WHAT I DESCRIBED HERE IS THE MOST COMMON USE CASE. IT'S NOT THE ONLY USE CASE. GOOGLE AND SOME OF THE OTHER COMPANIES HAVE ALSO BEEN DEVELOPING ALGORITHMS THAT RUN ON DEVICE. SO RATHER THAN HAVING TO SEND YOUR SPEECH SAMPLES OVER TO THE CLOUD, THEY -- ALL OF THE PROCESSING IS DONE LOCALLY. THOSE ARE CERTAINLY LESS COMMON BECAUSE, AS YOU MIGHT IMAGINE, AN ALGORITHM THAT'S BEEN CALIBRATED ON A MILLION HOURS OF SPEECH, IT RUNS VERY SLOWLY ON A CELL PHONE, AND SO IT WOULD DRAIN YOUR BATTERY IMMEDIATELY, AND SO MOST OF THE PROCESSING IS DONE ON THE CLOUD.

NOW, THIS CAN BE A PROBLEM. THERE'S BEEN A RECENT HEADLINE IN THE NEWS ABOUT THE VARIOUS COMPANIES, YOU KNOW, STORING THE SAMPLES AND LISTENING TO THEM AFTERWARDS. AGAIN, THIS IS DONE IN AN EFFORT TO IMPROVE THEIR ALGORITHMS, AND SO WHETHER THIS IS GOOD OR BAD REALLY IS A COMPLICATED QUESTION, AND I THINK WE'LL HAVE A FOLLOW-ON DISCUSSION ABOUT PRIVACY THAT I CAN EXPAND ON THIS A BIT MORE, BUT I JUST WANTED TO BRING IT TO EVERYONE'S ATTENTION NOW.

AND FINALLY, JUST ONE LAST SLIDE ON EVALUATING THE ASR ALGORITHMS, I THINK WE HAVE SOME FANTASTIC RESULTS WE'RE GOING TO SEE A BIT LATER. IN JUST A COUPLE OF WORDS HERE, WHAT'S IMPORTANT WITH THESE ALGORITHMS? THESE ARE CALIBRATED FOR SPECIFIC CONTEXT. SO BELIEVE IS IT OR NOT, GOOGLE HAS A SPEECH RECOGNITION ENGINE THAT'S SPECIFIC OR TRAINED ON YOUTUBE VIDEOS, ANOTHER ONE THAT'S BEEN TRAINED ON PHONE CALLS AND WORKS WELL IN CONVERSATIONAL PHONE SETTINGS, AND SO YOU HAVE DIFFERENT MODELS FOR DIFFERENT CONTENTS.

OF COURSE, HUMANS DON'T OPERATE LIKE THAT. WE USE THE SAME SET OF EARS. WE MODIFY THE WAY THAT WE LISTEN BASED ON CONTEXT. THAT'S CERTAINLY TRUE.

SO WHEN YOU EVALUATE THESE CONTEXT AND THE DESIGN OF THE TEST PARAMETERS, THE TEST SETTINGS, AS DIXIE MENTIONED IS ABSOLUTELY CRITICAL. WORD ACCURACY IS IMPORTANT; YOU WANT TO MAKE SURE THAT THE TRANSCRIPTS ARE CORRECTLY TRANSCRIBED, BUT SO IS LATENCY, VERY -- YOU KNOW, LONG, TIME PERIODS OF SILENCE KILL CONVERSATIONS, AND PRIVACY IS AN IMPORTANT PART OF THAT WE WILL DISCUSS LATER.

FINALLY I JUST REALIZED THAT THERE'S A REALLY UNFORTUNATE ERROR ON THE LAST BULLET, THAT SHOULD BE "CAN'T" NOT "CAN." THE MOST IMPORTANT METRIC IS ONE THAT CAN'T BE EASILY ~~KAWAUCHI~~

~~FIED.~~QUANTIFIED. SO THAT'S DEFINITELY CAN'T. IT TURNS OUT IT'S VERY DIFFICULT TO QUANTIFY HUMANS. AND THAT'S REALLY IT FOR NOW. WE'LL, YOU KNOW, DIVE INTO SOME OF THE OTHER QUESTIONS A BIT LATER. HOPEFULLY THAT WASN'T TOO LONG, DAVID.

>> DAVID BAHAR: . I REALLY APPRECIATE YOUR COMMENTS VISAR AND I WANT TO GET BACK TO A COMMENT YOU MADE ABOUT PRIVACY. YOU TALKED ABOUT CONVERSATION BEING STORED AND PERHAPS READ BY HUMANS LATER TO MEASURE FOR ~~AK ARE I SI.~~ ACCURACY. CAN YOU COMMENT FOR ABOUT THE RELATIONSHIP BETWEEN PRIVACY AND, SAY, HOW THIS MIGHT BE EMPLOYED IN RELAY SERVICES?

>> VISAR BERISHA: YEAH, THIS IS A REALLY IMPORTANT QUESTION. SO IF WE THINK ABOUT WHERE IMPROVEMENTS IN ASR CAME FROM, DURING THE LAST FEW YEARS, THEY'VE COME FROM THE AVAILABILITY OF DATA. THE FACT THAT PEOPLE SHARE THEIR SPEECH SAMPLES THAT ARE THEN TRANSCRIBED IN ORDER TO DEVELOP THESE MODELS. NOW, THAT SHARING COMES IN DIFFERENT FORMS. IT COMES FROM PEOPLE MAKING YOUTUBE VIDEOS, AND THEN THESE LARGE COMPANIES TRANSCRIBING THE RESULTING TEXT MANUALLY DEVELOPING THE ALGORITHMS. IT COMES IN SOME CASES FROM PEOPLE INTERACTING WITH THEIR DEVICES AND THEIR SPEECH SAMPLES BEING STORED.

SO IF WE THINK OF THESE VARIOUS CONSUMER APPLICATIONS, THIS AVAILABILITY OF DATA DURING THE LAST 15 YEARS OR SO HAS REALLY INCREASED EXPONENTIALLY, AND IT'S THAT, ALONG WITH SOME IMPROVEMENTS IN CONNOTATIONAL ABILITIES HAS RESULTED. PART OF THE ISSUE HERE IS THIS IS A VERY DIFFICULT PROBLEM TO BOUND, IN THE SENSE THERE'S A CLEAR BENEFIT TO COMPANIES AND CONSUMERS FROM HAVING ACCESS TO DATA ON A LARGE SCALE. IT CAN ENABLE NEW TECHNOLOGIES AND THEY WORK EXTREMELY, EXTREMELY WELL. SO WITH RESPECT TO CAPTIONS SERVICES SPECIFICALLY, TO BE HONEST IT'S QUITE DIFFICULT TO REALLY UNDERSTAND INTERNALLY WHAT'S HAPPENING WITH THE DATA BEHIND THESE COMPANIES.

THERE'S THESE VERY COMPLICATED TERMS AND CONDITIONS THAT MOST PEOPLE SCROLL DOWN AND HIT "ACCEPT" ON AS THEY'RE USING THE APPLICATIONS, AND I WOULD SAY THE LARGE-TECH COMPANIES HAVE LARGELY BEEN -- HAVEN'T BEEN VERY TRANSPARENT ABOUT HOW THEY HANDLE DATA. THE TRANSPARENCY COMES AFTER THERE'S SOME SORT OF LEAK OR THERE'S SOME NEWSPAPER STORY THAT COMES OUT ABOUT, YOU KNOW, THESE COMPANIES SENDING YOUR DATA OVER TO DIFFERENT HUMANS TO TRANSCRIBE SO THAT THEY CAN THEN FURTHER IMPROVE THEIR ALGORITHMS. SO

UNFORTUNATELY WHAT I GAVE YOU IS A NON-ANSWER ANSWER, AND PART OF THE REASON IS WE DON'T EXACTLY KNOW WHAT'S HAPPENING INTERNALLY. I KNOW THAT MOST OF THESE COMPANIES DECIDED THAT THEY'RE GOING TO PAUSE BECAUSE OF THESE NEWSPAPER STORIES THAT CAME OUT ON TRANSCRIPTION OF NEW DATA, BUT THEY VERY CAREFULLY SAID PAUSE AND DO NOT STOP. SO IT'S NOT CLEAR WHEN THAT WILL START AGAIN AND UNDER WHAT CONTEXT.

IF THERE'S A WAY, FOR EXAMPLE, FOR THAT TO BE GATED FOR CERTAIN APPLICATIONS, SO IT'S NOT USED FOR CAPTIONING SERVICES BUT MAYBE THEY COLLECT DATA IN OTHER CONTEXT, I DON'T KNOW, I THINK THE FUTURE WILL SORT OF TELL US THAT. BUT IT DOES PROVIDE -- I THINK IT'S A REAL IMPETUS FOR PEOPLE TO BE CURIOUS ABOUT WHAT'S HAPPENING WITH THEIR DATA AS THEY USE THESE DEVICES. IT'S VERY EASY TO USE THEM -- BECAUSE THEY'RE SO BENEFICIAL WITHOUT REALLY TAKING INTO ACCOUNT WHAT'S HAPPENING WITH THE DATA.

>> DAVID BAHAR: THANK YOU. WE'LL BE COMING BACK TO THAT PRIVACY DISCUSSION LATER ON, BUT BEFORE WE TURN OVER TO DIXIE AND CRE, I WANTED TO ASK YOU ONE MORE THING, IF YOU COULD SHARE SOME EXAMPLES OF WHERE ASR IS STILL STRUGGLING TO BE EFFECTIVE.

>> VISAR BERISHA: YES, YES, SO I THINK SOME OF THESE CONDITIONS, USERS OF THE TECHNOLOGY WILL BE, YOU KNOW, WELL-AWARE, NOISY ENVIRONMENTS, IT'S A STRUGGLE. ENVIRONMENTS WHERE THERE'S A LOT OF REVERBERATION, IT STRELS, SO REVERBERATION, IF YOU'RE IN CLOSE FACE AND YOUR AUDIO IS BOUNCING OFF THE WALLS SO THERE'S AN ECHO. ALTHOUGH IT'S GETTING BETTER EVERY DAY, MOSTLY BECAUSE OF THE AVAILABILITY OF DATA.

IT ALSO STRUGGLES WITH ATYPICAL ~~SPREECH~~ SPEECH, SO STRONGLY ACCENTED SPEECH, UNUSUALLY SLOW OR UNUSUALLY FAST SPEECH, THESE ARE ALL SCENARIOS UNDER WHICH THESE ASR ALGORITHMS STRUGGLE. BUT THERE'S ONE THING I WANT TO HIGHLIGHT, AND IT GOES BACK TO THE TESTING PARADIGM THAT'S USED. THE ALGORITHMS THAT ARE EMPLOYED ARE ONLY AS GOOD AS THE DATA THAT IS USED TO CALIBRATE THEM, TO TRAIN THEM.

OFTENTIMES, THE DATA USED TO TRAIN THESE ALGORITHMS ITSELF IS BIASED. WHAT DO I MEAN BY THAT? WELL, THERE HAVE BEEN STUDIES OUT THERE TO LOOK AT THE PERFORMANCE OF DIFFERENT SPEECH RECOGNITION ENGINES ON DIALECTS THAT ARE UNDER-REPRESENTED IN SPEECH. SO THIS TYPICALLY MEANS DIALECTS FROM INDIVIDUALS WITH LOWER ~~SOBER YES~~SOCIO ECONOMIC STATUS GROUPS. SO

THERE'S AN ~~ALGORITHM~~BIAS IN THE WAY THESE PERFORM AND ONE OF THE VARIABLES THERE IS SOCIOECONOMIC STATUS, WHO USES THE INTERNET AND WHO DOESN'T. IN SOME SENSE, BEING ABLE TO QUANTIFY THAT IN THE PARADIGM IS VERY IMPORTANT, ESPECIALLY AS WE TALK ABOUT DEMOCRATIZING THESE SERVICES TO EVERYBODY.

>> DAVID BAHAR: THANKS AGAIN. SO NOW, DIXIE AND CRE, I KNOW THAT OF YOU HAVE DONE SOME RESEARCH ~~CERTAINLIALLY~~ CERTAINLY IN TERMS OF HOW ASR WOULD APPLY TO TELEPHONE RELAY SERVICES, SO COULD YOU SHARE YOUR RESEARCH. ?

>> DIXIE ZIEGLER: YES, WE'D BE GLAD TO DO THAT. YOU CAN GO AHEAD AND PUT THE NEXT SLIDE THERE.

MAYBE ONE MORE. GREAT. SURE. ~~SDPLOOD~~ I'LL JUST GIVE YOU THE CLICKER.

>> DIXIE ZIEGLER: GREAT. SO ULTRATEC AND HAMILTON HAVE BEEN WORKING TOGETHER TO DO QUITE A BIT OF RESEARCH IN THIS AREA, AND CRE ALREADY MENTIONED THAT ULTRATEC HAS BEEN DOING THIS FOR A VERY, VERY LONG TIME, AND HAMILTON WAS THE FIRST CAPTIONED TELEPHONE PROVIDER, SO WE'VE HAD LOTS OF HISTORY IN WORKING ON ISSUES THAT ARE IMPORTANT TO CONSUMERS, WAYS TO IMPROVE QUALITY SERVICES, HOW CAN WE DELIVER THIS BETTER, HOW CAN WE BRING MORE VALUE TO THOSE WHO PAY FOR THESE SERVICES ARE ALL THINGS THAT WE'VE BEEN TALKING ABOUT FOR 20-PLUS YEARS, AND AS TECHNOLOGY AS CHANGED, CLEARLY THAT'S A PART OF OUR FOCUS AND WORK AS WELL, AND SO WANTING TO UNDERSTAND HOW TO APPLY AND BEST USE AUTOMATED SPEECH RECOGNITION INSIDE OF IPCCS HAS ALWAYS BEEN A FOCUS. I THINK ALL OF YOU KNOW THAT ASR HAS BEEN IN THE DELIVERY OF CAPTIONED TELEPHONE SERVICE FROM THE BEGINNING. THERE'S BEEN A USE OF AUTOMATED SPEAK RECOGNITION TOOLS, AND SO OBVIOUSLY, NOT NEW GROUND HERE, BUT CERTAINLY AS WE SAW THE CHART, IT'S REALLY THE AMOUNT OF DATA AND THE ABILITY TO USE THAT DATA HAS REALLY BEGUN TO GROW AND IMPROVE, AND SO WE'RE OBVIOUSLY WORKING TO UNDERSTAND ALL OF THOSE ISSUES.

SO AS PART OF THAT, WE WORKED WITH THE FCC TO DO SOME TRIALING OF ASR TECHNOLOGY AND APPRECIATED THE COMMISSION'S SUPPORT IN DOING SOME OF THIS RESEARCH, AND WE REALLY DID A THREE-PHASE TRIAL IN REALLY TRYING TO UNDERSTAND THE USE OF ASR AND IPCCS. THE FIRST PHASE OF OUR TRIAL WAS REALLY TO LOOK AT DIFFERENT ASR ENGINES, USED CONTENT RECORDED BY ~~VOL SUN TEARS.~~VOLUNTEERS. AFTER HAVING GONE THROUGH THAT PROCESS OF ANALYZING DIFFERENT ASR TOOLS, WE WENT ONTO FIELD

PERFORMANCE, TESTING IN THAT PHASE 2, AND COMPARED RESULTS BETWEEN THE DIFFERENT TYPES OF SERVICES AND HOW THEY'RE USED AND DELIVERED INSIDE OF IPCTS, BUT WHERE IT REALLY GOT FUN AND EXCITING IS IN PHASE 3 AND THAT'S WHERE WE'RE GOING TO SPEND THE MAJORITY OF OUR TIME TALKING WITH YOU TODAY. AND THAT'S WHERE WE HAD VOLUNTEERS WHO WERE ACTUALLY USING THE PHONE, USING THE SERVICE THE WAY THEY ALWAYS HAD SO THEY WERE ALREADY CAPTION TELEPHONE USERS, HAD THE EQUIPMENT IN HIR HOME, USING THE SERVICE FOR QUITE SOME TIME. AND THESE FOLKS TRIED TO USE THE SERVICE IN A NEW WAY, DICHT WAY, SO WE HAD INDIVIDUALS FROM OUR TEAMS TRAIN THESE VOLUNTEERS, THESE TESTERS, TO UNDERSTAND HOW TO USE THE SYSTEM, HOW TO USE THE PHONE, AND THEN THEY WERE ABLE TO SCORE -- WE WERE THEN ABLE TO TAKE SOME INFORMATION FROM THESE USERS.

SO AS I MENTIONED WE HAD VOLUNTEERS, THEY WERE USING THE EQUIPMENT THAT THEY WERE ALREADY FAMILIAR WITH. VOLUNTEERS WERE SURVEYED AFTER EACH OF THE CALLS THAT THEY PLACED. THEY WERE USING THE PHONE AS THEY NORMALLY WOULD, SO THEY WERE OWN CONVERSATIONS ~~SHG~~ USING THE PHONE IN THEIR OWN ENVIRONMENT FOR WHATEVER KINDS OF CALLS THEY MADE. WE THEN MEASURED THE PERFORMANCE AND USEFULNESS OF THOSE CAPTIONS FROM THE TRIALISTS. WE DID PAY THESE TRIALISTS AND WE DID SO BECAUSE IN SUCH A MANNER AS PART OF A GOOD RESEARCH AND STUDY TENETS SO WE COULD GET THE BEST FEEDBACK WE COULD, OBVIOUSLY, WE WANTED TO JUST RECOGNIZE THOSE USERS AND THEIR TIME FOR GIVING US THEIR INPUT. AND WE DID NOT BILL ANY OF THESE CALLS TO THE FUND. AGAIN WE WERE GRATEFUL FOR THE FCC TO GIVE US THE ABILITY TO DO THIS TRIAL. ANYTHING ELSE YOU WANT TO SAY? ANYTHING YOU WANT TO ADD THERE OR SHOULD I ~~MUF~~ MOVE ON?

>> CRE ENGELKE: SO I GET TO PUT ON THE BIG NERD HAD AND TALK A LITTLE BIT MORE ABOUT HOW WE DID YOU GO INTO THE DETAILS OF THIS STUDY. I THINK THE ONLY THINGS THAT I WOULD ADD THERE, SO, ONE, WE GAVE PEOPLE IN A DOUBLE BLIND WAY EITHER ASR CAPTIONS OR CA-BASED CAPTIONS AND WHAT THIS DID IS GAVE US A RANGE ACROSS ACCURACY AND DELAY. SO WE HAD CALLS THAT WERE VERY ACCURATE, BUT QUITE DELAYED. WE HAD CALLS THAT WERE VERY ACCURATE AND VERY FAST. WE HAD CALLS QUITE DELAYED AND QUITE INACCURATE AND SO ON. I MEAN, EVERY COMBINATION OF THOSE TWO METRICS THAT YOU CAN IMAGINE, AND WE ALSO SURVEYED THE USERS, AND I'LL TALK A LITTLE BIT MORE ABOUT THIS, ACTUALLY, RIGHT NOW.

WE SURVEYED THEM ACROSS A SERIES OF THREE BASIC AREAS. SO WE ACTUALLY UNDERWENT THE FAIRLY MASSIVE STUDY AND I WON'T DRAG YOU THROUGH IT, BECAUSE, ONE, MY NERD HAT MAY GET TOO BIG FOR THIS ROOM. WE TOOK ABOUT 10,000 OR 20,000 USER COMMENTS, USER GROUPS, FOCUS GROUPS AND SO FORTH AND WE LOOKED AT WHAT ARE THOSE THINGS THAT PROP UP FUNCTIONAL EQUIVALENCY, AND FUNCTIONAL EQUIVALENCY AS YOU ALL KNOW IS THE BUZZ WORD MANDATE, IT'S WHAT MAKES RELAY SERVICE RELAY SERVICE IN THE UNITED STATES. IT IS THE CORE TO WHAT USERS EXPECT, DEMAND, AND WHAT THEY HAVE THE RIGHT TO.

SO TRADITIONALLY FUNCTIONAL EQUIVALENCY HAS BEEN UNDERSTOOD IN TERMS OF SPEED AND ACCURACY. AND THOSE ARE GOOD THINGS TO MEASURE, BECAUSE THEY'RE REALLY EASY TO MEASURE. AS VISAR SAID THE HARDEST THING TO QUANTIFY IS ALL THOSE THINGS THAT PRECIPITATE FROM THAT. ALL THOSE THINGS THAT MAKE THIS AN EQUIVALENT SERVICE. ALL THOSE THINGS THAT USERS ACTUALLY RELY ON. SO WE WENT THROUGH THIS MASSIVE STUDY, AND THE AREAS THAT WE FOUND, PSYCHOLOGISTS THAT WE EMPLOYED AND STATISTICIANS WE EMPLOYED USED THE TERM LATENT VARIABLES, THE THINGS ~~LADDER~~ HARDER TO MEASURE IN THE WORLD BUT YOU CAN GET AT IN CERTAIN WAYS. AND PEOPLE TALK ABOUT THREE THINGS, THEY TALKED ABOUT THE EXPERIENCE OF BEING ON A PHONE CALL. THIS IS THE BEING ~~BEING~~ PRESENT WITH SOMEONE, THE TURN-TAKING THE ~~REPERTOIRE,~~ REPARTEE, THIS THE SECOND IS THE CONFIDENCE, THE FAITH IS THE TECHNOLOGY, THE FACT THAT THIS IS GOING TO SUPPORT THEM, WE ASKED QUESTIONS LIKE: DO YOU TRUST THIS EQUIPMENT? DO YOU TRUST THESE CAPTIONS AND SO FORTH.

AND THEN THE THIRD WAS EMPOWERMENT. IT DRIVES AT QUESTIONS LIKE: WOULD YOU FEEL SAFE USING THIS IN AN EMERGENCY? DO YOU FEEL YOU COULD USE THIS ON IMPORTANT BUSINESS CALLS? TALKED ABOUT NOTIONS OF INDEPENDENCE. AND THESE WERE THE THREE AREAS THAT THE USERS TOLD US OVER THESE X MANY 10S OF THOUSANDS OF DATA POINTS THAT THAT'S WHAT MADE FUNCTIONAL EQUIVALENCY FOR THEM; THAT'S WHAT THEY NEEDED THE PHONE FOR ON AND FORTH.

THOSE ARE THE AREAS THAT WE SET OUT AND SURVEYED AT THE END OF EVERY CALL. AND THEN WHAT WHAT WE DID AS DIXIE MENTIONED WE HAD NUMBERS, WE HAD DATA THAT WE HAD MEASURED ON EVERY CALL, SO WE COULD TELL YOU HOW DELAYED THE CAPTIONS WERE ON THAT CALL, AND WE COULD TELL YOU HOW ACCURATE THAT CALL WAS. AND WE COULD CROSS-REFERENCE HOW PEOPLE EVALUATED THESE THREE AREAS OF EXPERIENCE, CONFIDENCE AND EMPOWERMENT, AND LOOK AT THE RELATIONSHIP BETWEEN SPEED AND ACCURACY

AND EXPERIENCE, CONFIDENCE, AND EMPOWERMENT. SOUNDS SO SIMPLE WHEN I SAY IT LIKE THAT, AND YET I CAN TELL YOU THIS TOOK ME YEARS, WHICH FELT LIKE DECADES. . .

I SHOULD ALSO POINT OUT THAT WE DID THIS WITHOUT RECORDING ANY CALL ~~CONSENT.~~ CONTENT. SO WE ACTUALLY DEVELOPED TOOLS TO ALLOW US TO ASSESS THE ACCURACY IN REALTIME SO WE DIDN'T HOLD ANY CONTENT. SO WHEN PEOPLE HUNG UP. . .

SO THIS IS WHAT WE FOUND AND IF I'M RUNNING OVERTIME, DAVID PLEASE KICK THE BACK OF MY CHAIR.

>> DAVID BAHAR: SO FAR SO GOOD BUT I WILL.

>> ~~VISAR BERISHA~~ CRE ENGELKE: AND IF I GET TOO NERDY AND BORING, PUT YOUR HEAD DOWN AND I'LL KNOW TO DANCE OR SOMETHING, MORE INTERESTING.

BUT THIS IS WHAT WE FOUND. WHAT YOU SEE ON THE SCREEN HERE --

>> DAVID BAHAR: ANY VOLUNTEERS TO COME UP AND DANCE?

>> ~~VISAR BERISHA~~ CRE ENGELKE: OKAY. (LAUGHTER) WHAT YOU SEE HERE ARE THREE DIFFERENT SETS OF RELATIONSHIPS BETWEEN THE DELAY IN RED AND THE ACCURACY IN BLUE ACROSS THE THREE DIFFERENT MEASURES THERE OF EXPERIENCE, CONFIDENCE, AND EMPOWERMENT. AND WHAT THIS TELLS YOU IS HOW MUCH OF THE VARIATION, HOW WELL DID THE ACCURACY OR THE DELAY PREDICT HOW PEOPLE WERE GOING TO ANSWER QUESTIONS ABOUT THAT VARIABLE, ABOUT -- SO IN THE FIRST CASE, EXPERIENCE.

AND WHAT WE FOUND -- AND THIS WAS ACTUALLY QUITE SURPRISING TO ME. I MEAN, THIS WAS THE ONE THAT I ACTUALLY GOT SHOCKED BY. IN EVERY STUDY, AND VISAR CAN HOPEFULLY BACK ME UP HERE, YOU ALWAYS FIND SOMETHING THAT'S GOING TO SURPRISE YOU, AND THIS IS THE ONE THAT MADE ME GO, WHOA. WHAT WE FOUND IN EXPERIENCE WAS THAT ACCURACY WAS TWO TO THREE TIMES MORE IMPORTANT THAN DELAY IN HOW PEOPLE EXPERIENCED THAT CALL, IN HOW PEOPLE EXPERIENCED THAT SENSE OF CLOSENESS AND RELATIONSHIP WITH THE OTHER PERSON ON THE CALL.

AND WHY DID IT SURPRISE ME? WELL, AGAIN, BECAUSE I EXPECTED ALL OF THIS TO BE ABOUT, HEY, I'M WITH THIS OTHER PERSON, SO RAPID TURN-TAKING IS IMPORTANT, AND LET'S GET CAPTIONS OUT RIGHT AWAY AND LATENCY SIMP SO FORTH. AND THAT'S NOT WHAT PEOPLE SAID. WHEN WE INTERVIEWED PEOPLE, WHAT THEY TOLD US IS, NO, I

NEED THE GET THE CAPTIONS RIGHT AWAY, BUT I NEED THE CAPTIONS TO BE RIGHT, RIGHT AWAY. BECAUSE IF I GET A WRONG CAPTION AND THEN THERE'S A CORRECTION OR IF I HAVE TO READ IT AND IT DOESN'T QUITE MAKE SENSE, IT THROWS THE WHOLE THING OFF. AND JUST A LITTLE FORESHADOWING, THIS IS A WHOLE NEW RESEARCH PROJECT OR BECAME A RESEARCH PROJECT FOR US. I WAS AT HLAA IN MINNEAPOLIS, AND DAVID SCHMIDT WAS TALKING ABOUT THIS, SAID, MAN THIS IS REALLY HARD, BECAUSE YOU DO RESEARCH AND YOU FIND 20 NEW THINGS YOU HAVE TO LOOK AT.

SO PEOPLE WANT IT QUICKLY, BUT IT HAS TO BE RIGHT. THEY WANT IT FAST, BUT IF IT'S WRONG, IT CAN THROW THE ENTIRE SEQUENCE OF CONVERSATION INTO CHAOS. I REALLY WASN'T SURPRISED THAT ACCURACY WAS IN THIS CASE FIVE TIMES FOR POWERFUL THAN DELAY IN PREDICTING HOW PEOPLE WERE GOING TO RANK THEIR CONFIDENCE MEASURES, BECAUSE AS ONE PERSON SAID, I'VE LEARNED NOT TO TRUST MY EARS. I'M USING A CAPTION TELEPHONE BECAUSE I DON'T TRUST THAT I'M FAITHFULLY UNDERSTANDING WHAT THE OTHER PERSON IS TELLING ME.

SO THE QUESTION HERE IS, DO I TRUST MY EYES? DO I TRUST THESE CAPTIONS TO ACCURATELY REPRESENT WHAT'S BEING SAID? SO OBVIOUSLY CONFIDENCE HINGES ON ACCURACY.

IF YOU CAN'T READ IT THE FARTHEST LEFT IS USER EXPERIENCE, AND THE CENTER ONE IS CONFIDENCE, AND THEN THE FURTHEST RIGHT IS EMPOWERMENT, AND THAT'S WAS FINAL ONE. ERIC, DID YOU HAVE A. . .

>> YEAH, IF I MAY.

>> HOLLY BISE: ERIC, WAIT.

>> ERIC JESCHKE: JUST, WHAT YOU SAID IS EXACTLY THE POINT. IF YOU'VE EVER PLAYED BINGO, HE OR SHE HAS A RHYTHM. THAT RHYTHM CAN BE FAST OR SLOW, BUT BOY, IT BETTER BE RIGHT. IF YOU THINK ABOUT IT, IF THE CALLER MAKES A MISTAKE, THAT MESSES EVERYBODY, AND EVERYBODY IS, YOU KNOW, IN A PROBLEM. SO WHAT YOU SAID, I WAS JUST SAYING, THAT'S A GOOD EXAMPLE. IF YOU EVER PLAYED BINGO, GOT IT?

>> CRE ENGELKE: THAT'S A SPECTA SPECTACULAR EXAMPLE. THANK YOU. I LOVE THAT. I'VE BEEN USING THE ONE ABOUT IF YOU'RE DRIVINGS HOME, YOU WANT TO GET HOME QUICKLY, BUT IF YOU GO TO THE WRONG HOME, THEN IT DOESN'T REALLY COUNT. BUT, THANK YOU.

MY LAST POINT WAS ONLY IN TERMS OF EMPOWERMENT, WHAT WE SAW WAS THAT ACCURACY HERE WAS ABOUT THREE TIMES MORE POWERFUL THAN DELAY IN PREDICTING HOW PEOPLE WERE GOING TO RESPOND TO THE EMPOWERMENT QUESTIONS. QUESTIONS ABOUT ~~CONTINUOUSING~~ CONTINUING SENSES OF INDEPENDENCE, SENSES OF SAFETY, THE ABILITY TO CONTINUE BEING WHATEVER ROLE, FATHER, BOYFRIEND, HUSBAND, FRIEND, WHATEVER, HOW DO YOU GO ABOUT DOING THAT ROLE DEPENDS ON THIS ACCURACY.

AND, AGAIN, I THINK THE COMMENT THAT STUCK OUT FOR ME THE MOST STRONGLY WHEN WE WERE INTERVIEWING PEOPLE AT THE END OF THIS, AND BY THE WAY, I DON'T KNOW IF ANYONE IN THIS ROOM KNOWS THIS, BUT THERE'S ALTS BIT OF A STIGMA AROUND HEARING LOSS, AND IF YOU HAVE TO SAY "WHAT" ABOUT A HUNDRED TIMES IN A PHONE CALL, PEOPLE DON'T WANT TO TALK TO YOU ANYMORE. AND THAT'S WHAT THIS GUY WOULD SAY. HE SAID I CAN'T GO ABOUT DOING MY WORK IF MY CLIENTS HAVE TO REPEAT THEMSELVES. THEY WANT TO TALK TO SOMEONE ELSE. THAT WAS THE ONE THAT ACCURACY MADE SENSE.

WELL, AND I'M ALMOST DONE, SO I'LL JUMP OFF IN A SECOND HERE. BUT LIKE ALL GOOD ACADEMICS, WHAT WE FOUND, WE FOUND ALL OF THESE THINGS AND WE THOUGHT, OKAY, IF YOU FIND SOMETHING SOMEWHERE, YOU BETTER VERIFY IT. YOU'VE GOT TO GO FIND IT SOMEWHERE ELSE. AND SO WE LOOKED AT ALL THE DATA THAT WE HAD, OR A LOT OF THE DATA WE HAD FROM THESE VOLUNTEERS, AND AT ONE POINT IN THE CALL WE ASKED THEM OR AFTER THE CALL WE ASKED THEM TO ESTIMATE THE ACCURACY AND THE DELAY ON EACH OF THEIR CALLS, AND THEN WE ASKED THEM THESE QUESTIONS. AND WHAT WE FOUND WAS THAT NOT ONLY DID THEY PERCEIVE ACCURACY TO BE 2 TO 5 TIMES FOR POWERFUL THAN DELAY, BUT FOR THEM IT WAS ACTUALLY CLOSER FIVE TO TEN TIMES MORE POWERFUL THAN DELAY. SO WHEN SOMEONE ESTIMATED THE ACCURACY USING THEIR ESTIMATE, THEY WERE LIKE NO EITHER IT WAS ACCURATE AND IT HELPED ME A LOT AND ALL OF THESE THINGS TODAY TOGETHER OR IT WAS INACCURATE AND EVERYTHING ELSE DROPPED.

SO IF I WAS TO PLACE A BET ON ANY GIVEN CALL ON HOW SOMEONE WOULD RESPOND TO THESE THINGS, BY LOOKING AT THIS, ALL I REALLY NEEDED TO KNOW WAS HOW ACCURATE IT WAS AND HOW FAITHFUL THE REPRESENTATION.

NOW, WE LEARNED A LOT DOING THIS STUDY AND ONE OF THE THINGS WE LEARNED WAS HOW TO MEASURE ACCURACY, OR HOW TO MEASURE ACCURACY AND WHAT TYPES OF PITFALLS DIFFERENT TYPES OF ERRORS CREATE. SO IT'S NOT JUST WRONG OR RIGHT. IT'S

DIFFERENT TYPES OF WRONG, AND DIFFERENT TYPES OF RIGHT. AND THOSE HAVE DIFFERENT TYPES OF IMPACT ON HOW PEOPLE PROGRESS BOTH IN UNDERSTANDING AND IN THE FLOW OF THE CONVERSATION. THAT'S ALL FOR ME. MAYBE PICK THAT UP LATER. OH, SORRY, I HAD ONE MORE POINT.

WE FOUND SOME WEIRD STUFF. ACCURACY AND DELAY, THE PERCEPTIONS OF ACCURACY AND DELAY WERE NOT INDEPENDENT. SO PEOPLE DIDN'T EXPERIENCE ACCURACY AND DELAY AS TWO SEPARATE MEASURES, BUT THEY ACTUALLY BLED INTO ONE ANOTHER. SO HOW ACCURATE SOMETHING WAS WOULD ACTUALLY CHANGE HOW QUICKLY PEOPLE PROCESSED THAT INFORMATION AND HOW MUCH THEY THOUGHT, HOW MUCH THEY THOUGHT OF THE DELAY. SO, THERE'S A LOT OF THINGS HERE WHERE WE THOUGHT WE WERE DEALING WITH INDEPENDENT VARIABLES, YOU KNOW, ACCURACY AND DELAY, BUT IN FACT THEY'RE NOT, THEY'RE DEPENDENT AND THEREFORE NETWORKED. SO THANK YOU.

>> DAVID BAHAR: THANKS. THAT WAS AWESOME. I ~~THINK~~ WANT TO THANK CRE AND DIXIE FOR THAT. AND DOES ANYONE WANT TO COVER MY NEXT QUESTION? I THINK YOU ALREADY COVERED IT. I CAN GIVE YOU 5 MINUTES BACK. I HAVE MORE QUESTIONS TO ASK, BUT SINCE WE STARTED EARLY, I P DIDN'T KNOW HOW MUCH TIME WE HAVE AT THE END FOR YOU ALL TO ASK YOUR QUESTIONS, AND I WANT TO MAKE SURE WE CAN DO THAT.

SINCE WE ALREADY COVERED MY QUESTIONS, THAT ~~GIFZ~~ GIVES US BACK A LOT OF TIME FOR ANY BURNING QUESTIONS THAT YOU MAY HAVE, AND IF YOU HAVE SOME, WE'RE GOING TO BE GETTING THERE PRETTY SOON, SO DON'T WORRY.

>> CRE ENGELKE: I WOULD LOVE TO ACTUALLY ANSWER SOME OF THIS QUESTION, BECAUSE I DON'T WANT TO --

>> DAVID BAHAR: PLEASE DO.

>> CRE ENGELKE: SO I DON'T WANT TO GO ON RECORD AS SAYING DELAY DOESN'T MATTER OR ANYTHING FOOLISH LIKE THAT. I THINK THE ANSWER TO THIS REQUEST IS I AGREE THAT ACCURACY AND DELAY BOTH MATTER. SO I WASN'T TRYING TO SUGGEST THAT.

BUT ONE OF THE THINGS THAT'S COMING OUT IN SOME OF OUR NEWEST RESEARCH IS THERE'S DIFFERENT TYPES OF ERRORS AND THEY CAUSE DIFFERENT TYPES OF DISRUPTION TO THE CONVERSATION. AND THOSE DISRUPTIONS CAN COME IN A NUMBER OF

WAYS. SO LET ME JUST GIVE YOU A COUPLE OF EXAMPLES. JUST TWO EXAMPLES OF DIFFERENT TYPES OF ERRORS.

WE HAVE WHAT'S CALLED A VISIBLE ERROR AND INVISIBLE ~~OR.~~ ERROR A VISIBLE ERROR IS ONE THAT MAKES YOU GO HUH, SO IF I SAY I'M GOING TO THE STORE TODAY, AND IT IF THE CAPTIONS HOME OUT I'M HIPPOPOTAMUS TO THE STORE TODAY, YOU'RE GOING TO GO, HUH? THAT DELAY HAS AN IMPACT ON HOW THAT FLOW GOES AND HOW THAT IMPACT GOES, YOU SAY WHAT? AND NOW WE'RE IN A REPAIR SEQUENCE WHERE YOU'RE REPAIRING WHAT I'VE SAID.

THE OTHER TYPE OF ERROR IS AN INVISIBLE ERROR, AND I APOLOGIZE IN ADVANCE TO THE CART, MY PHONE ~~FLB~~ NUMBER IS 3105695. WELL, THANK YOU TRYING. HOW ABOUT MY PHONE NUMBER IS ~~###-###-####.~~ THAT'S ACTUALLY MY NUMBER. SORRY. WELL, WHAT IF THE ~~CAPTIONED~~ CAPTIONS HAD COME OUT MY NUMBER IS ~~###-###-####~~ YOU'RE NOT EVEN GOING TO GET SOMEONE WHO KNOWS ME. IN FACT YOU'RE GOING TO GET THE HAMPTON ~~INSOMEWHERE~~IN INN SOMEWHERE ON THE EAST COAST, AND I KNOW THIS BECAUSE SOMEHOW PEOPLE REVERSE THAT 1 AND ~~SDPLER~~ THE 0 ALL THE TIME AND I'M CONSTANTLY MAKING RESERVATIONS FOR PEOPLE AT MY HOUSE NOW. BUT THE POINT IS, THAT IS A TOTALLY ~~AN~~ INVISIBLE ERROR. THERE'S NO REASON FOR YOU TO ASK FOR CLARIFICATION, BECAUSE IT'S RIGHT THERE. IT'S 10 DIGITS, IT CONFORMS TO THE STRUCTURE OF AN EXPECTATION OF HOW YOU WOULD EXPECT THOSE NUMBERS TO COME OUT. WHY WOULD YOU ASK SOMEONE, WHY WOULD YOU SAY "WHAT?" SO THE REPAIR THAT HAPPENS, IF IT HAPPENS, HAPPENS MUCH FURTHER DOWN IN THE ~~KVRS~~ CONVERSATION, AND IT TENDS TO BE A DIFFERENT TYPE OF REPAIR. ALSO, INTERESTINGLY, IT TENDS TO BE THAT THE PERSON WHO GETS BLAMED FOR THAT ERROR, FOR THAT INVISIBLE ERROR IS THE PERSON WHO SPOKE IT, WHEREAS THE PERSON WHO SAYS "WHAT?" GETS BLAMED FOR NOT HEARING IT. SO WE'RE LOOKING AT ALL OF THESE DIFFERENT WAYS OF CLASSIFYING ERRORS AND HOW THEY IMPACT USERS, AND FROM THAT, HOW BEST TO UNDERSTAND NOT JUST THE TIMING OF WHEN A WORD SHOWED UP, BUT HOW THAT TIME IMPACTS THE OVERALL FLOW OF THE CONVERSATION FOR THE USER.

SO THAT'S KIND OF WHERE I WANTED TO GO. I THOUGHT THAT REALLY DESERVED TO BE UNDERSTOOD; THE USER'S EXPERIENCE HAS TO BE.

>> DIXIE ZIEGLER: THAT'S ONE OF THE THINGS THAT THE INDUSTRY GROUP IS REALLY TRYING TO ANALYZE AND BREAK DOWN AND DISCUSS AND AGREE TO THESE KINDS OF WHETHER IT'S A DEFINITION, OR WHAT WE SHOULD BE MEASURING OR HOW WE SHOULD BE ~~MAEFRING~~ MEASURING AND MANY OF YOU PROBABLY SAW THAT THE

INDUSTRY GROUP LAST FALL HAD ACTUALLY PUT SOME COMMENTS IN TO THE FCC, WE SET ~~SIMP~~ DID AN EX PARTE, AND WE'RE ~~PARKING~~ PLANNING ON ANOTHER, AND IT'S VERY, VERY CLOSE TO GIVING YOU AN UPDATE ON WHAT THE INDUSTRY HAS BEEN DOING OVER THE LAST YEAR, AND THERE'S BEEN SOME REALLY GREAT WORK DONE. A LOT OF DISCUSSION ABOUT WHAT CRE ~~CRE~~ WAS JUST TALKING ABOUT AND ALL OF THE PROVIDERS CONTRIBUTED TO THAT AND HOW DO WE BEST BUILDING A TESTING SYSTEM AND ANALYZE OUR DATA AS A RESULT OF THAT, MAKING SURE THAT WE HAVE THE RIGHT PARAMETERS, THE RIGHT TESTING METRICS, THE RIGHT DEFINITIONS. HOW DO WE MAKE SURE WE PULL IN A USER EXPERIENCE INTO THAT? HOW CAN WE USE ALL OF THAT DATA TO AS I MENTIONED AT THE BEGINNING, FORMULATE THE RIGHT QUALITY METRICS. SO THE INDUSTRY GROUP HAS SPENT A TREMENDOUS AMOUNT OF TIME TALKING ABOUT WHAT WE CALL A SCORECARD AND HOW THE DIFFERENT THINGS THAT WE SHOULD MEASURE AND TALKED ABOUT TESTING AND SAMPLE SIZE, AND WHO IS REPRESENTED. AND WE TALKED A LITTLE BIT ABOUT BIAS IN SOME OF THE SAMPLING THAT ASR ENGINES HAVE USED. WELL, HOW DO WE AVOID THAT? BECAUSE WE RECOGNIZE THAT TO BE EQUIVALENT HAS TO WORK FOR ALL, IRN THAT SOUNDS LIKE BAR, AND IT SOUNDS LIKE A BIG GOAL, BUT IT HAS TO WORK FOR ALL, SO HOW DO WE MAKE SURE WE ARE PUTTING IN THE RIGHT TESTING MEASUREMENTS AND METRICS TO MAKE THAT WORK FOR ALL. SO THE INDUSTRY GROUP HAS CONTINUED TO PUT MEAT ON THE BONE HERE. WE'RE WORKING TO ~~FIEBD~~ FIND A THIRD PARTY TO DO THE MEASUREMENTS, TO DO THE DATA ANALYSIS AND BRING THAT IN A WAY FORWARD THAT WE CAN CONTINUE TO SHARE, AND CONTINUE THE DIALOGUE AROUND QUALITY METRICS, SO OUR INDUSTRY GROUP IS EXCITED ABOUT THAT AND HOW WE'RE TRYING TO ENSURE WE AVOID THE BIAS, GET TO THE RIGHT ANSWERS AND REALLY MOVE FORWARD ON OUR QUALITY METRIC ROAD MAP.

>> DAVID BAHAR: THANK YOU, BOTH. I HOPE THAT YOU'VE BEEN THINKING OF QUESTIONS YOU'D LIKE TO ASK. AND I'D LIKE TO MAYBE ASK ANOTHER COUPLE OF QUESTIONS, MAYBE THREE MINUTES EACH TO RESPOND, AND THEN OPEN IT UP TO QUESTIONS YOU MAY HAVE, AND WE CAN LINE UP AS YOU DID BEFORE AT THE ~~MIKE RE PHONES~~ MICROPHONES OR IF YOU'RE SIGNING, YOU CAN COME UP TO THE FRONT. AND WE PLAN TO HAVE ABOUT 15 MINUTES FOR THAT. SO MY QUESTION FOR VISAR, IN YOUR PRESENTATION, THERE WAS INFORMATION ABOUT CORTANA, ALEXA, SIRI AND THE LIKE AND GOOGLE AND HOW SOME OF THEM ARE MAINTAINING CONVERSATIONS SO WITH ASR, HOW CAN CONSUMERS BE CONFIDENT IF THEIR CONVERSATIONS ARE BEING STORED SOMEWHERE OR NOT.

>> VISAR BERISHA: WHAT A GREAT QUESTION. SO ONE OF THE BIGGEST SCAMS ON THE INTERNET IS THAT TERMS AND CONDITIONS STAPLE. NO ONE READS THOSE THINGS. PEOPLE TYPICALLY SCROLL ALL THE WAY TO THE BOTTOM AND THEY HIT "ACCEPT" AND MOVE ON. SO HOW CAN THEY KNOW? THE INFORMATION IS THERE, BELIEVE IT OR NOT, FOR A LOT OF THESE QUESTIONS THEY PUT IT THERE, EITHER DIRECTLY OR INDIRECTLY. THAT HAS TO CHANGE SOMEHOW. THOSE STATEMENTS AND THE IMPORTANT INFORMATION ABOUT DATA, WHERE IT'S STORED AND HOW IT'S STORED, FOR HOW LONG IT'S STORED, HAS TO COME UP FRONT. THERE'S SOME MOVEMENT AROUND THIS. I'VE SEEN THIS AROUND WITH SOME COMPANIES, AND THIS IS ESPECIALLY TRUE IN HEALTHCARE, WHERE THEY WILL TAKE THE TERMS AND CONDITION STATEMENT AND SIMPLIFY IT IN EASY TO UNDERSTAND BULLET POINTS AND NOT, YOU KNOW, LAWYER-SPEAK. SO UNTIL THAT HAPPENS, AND UNTIL THAT BECOMES A STANDARD THROUGH THE INDUSTRY, REALLY, THE ONLY WAY THAT THEY CAN KNOW IS THROUGH VISIBILITY, THROUGH GOING THROUGH AND MAKING SURE THAT THEY SORT OF UNDERSTAND THE TERMS AND CONDITIONS ASSOCIATED WITH THE PARTICULAR PROVIDER OF THE SERVICE.

YOU KNOW, IT'S INTERESTING, AFTER THESE NEWSPAPER ARTICLES CAME OUT ABOUT THE FACT THAT THESE COMPANIES WERE STORING THE SPEECH, SO GOOGLE AND APPLE AND THE LIKE IMMEDIATELY SAID THAT WE'RE GOING TO PAUSE THESE PROGRAMS. MICROSOFT SIMPLY UPDATED THEIR TERMS AND CONDITIONS AND ADDED IT INTO THEIR THAT SAYS" WE MAY BE RECORDING YOUR SKYPE CALLS. AND THAT'S IT. SO, YOU KNOW, UNLESS PEOPLE SORT OF TAKE INITIATIVE, UNLESS THIS BECOMES A BIT MORE TRANSPARENT, I DON'T KNOW THAT THERE'S AN EASY ANSWER TO THAT QUESTION.

>> DAVID BAHAR: THANKS VISAR. NOW I HAVE A QUESTION FOR DIXIE AND CRE. CURRENTLY CAPTION TELEPHONE SERVICES USE SPEECH RECOGNITION WITH THE ASSISTANCE OF A CALL SYSTEM TO CLEAN UP THE MESSAGE BEFORE IT GOES OUT TO THE CONSUMER. HAVE YOU FOUND THAT CONSUMERS CAN EASILY SWITCH BETWEEN TRADITIONAL CAPTION TELEPHONE IT SERVICE AND ASR SERVICE AND IF YOU WERE ABLE TO DO THAT, WHAT DO YOU THINK YOU WOULD FIND?

>> CRE ENGELKE: THAT'S A GREAT QUESTION. AS I SAID, WE'VE DONE A NUMBER OF DIFFERENT RESEARCH PROJECTS, ALL PART OF -- EXTENDED RESEARCH ON THIS. SO THE ANSWER TO THE QUESTION IS REALLY A TWO-PART ANSWER. IS IT TECHNICALLY POSSIBLE TO ENABLE PEOPLE TO SWITCH FROM A CA-BASED CAPTIONS SVZ TO AN ASR SERVICE AND BACK AND FORTH BY HITTING A BUTTON? AND THE ANSWER IS YES. YES, TECHNICALLY IT'S QUITE POSSIBLE, AND IN FACT WE DID IT FOR ONE ~~FR~~ OF OUR TESTS. THE SECOND QUESTION, THEN,

IS, IS IT PRACTICAL? IS THIS A REASONABLE APPROACH? DOES IT ENABLE THE INTERACTION AND QUALITY AND AGENCY OVER CAPTIONS THAT A USER MIGHT WANT SNR AND OUR FINDINGS ARE THAT IT IS NOT, NOT IN ANY WAY A GOOD WAY OF APPROACHING THAT.

SO WE HAD USERS, QUITE A FEW USERS WHO WE BASICALLY -- WE WENT IN, WE HAD SOMEONE GO TO THEIR HOME AND TRAIN THEM ON A SPECIAL PIECE OF EQUIPMENT THAT WOULD ALLOW THEM TO SWITCH FROM ASR TO CA-BASED CAPTIONS BY HITTING A BUTTON. WE TRAINED THEM, EVERYTHING WENT AS PLANNED IN THAT REGARD. WE LEFT THEM TO THEIR OWN DEVICES FOR A WEEK OR SO TO USE IT BY THEMSELVES IN A WAY WE TALKED ABOUT EARLIER. AND WHAT WE FOUND IS BY HITTING THAT BUTTON -- FIRST OF ALL, BEING RESPONSIBLE FOR THE QUALITY OF CAPTIONING AND RECOGNIZING HOW USEFUL OUR ACCURATE YOUR CAPTIONS ARE TAKES ANOTHER LAYER OF AWARENESS THAT IS JUST SO FAR ABOVE AND OUTSIDE OF WHAT PEOPLE TYPICALLY EXPERIENCE THAT FOR THE MOST PART PEOPLE JUST FORGOT IT WAS THERE. THEY JUST SIMPLY DIDN'T USE THAT SERVICE AT ALL. THEY HAD CALLS THAT SUFFERED TERRIBLE ACCURACY, THE KIND OF ACCURACY THAT YOU WOULD NEVER ACCEPT FROM ANYTHING, AND THEY DIDN'T HIT THE BUTTON.

ON THE OTHER SIDE -- AND QUITE FRANKLY, JUST, YOU KNOW, TO THINK ABOUT IT, THE LEVEL OF SORT OF COGNITIVE LOAD THIS PUTS THE USER UNDER, IT DOESN'T QUALIFY UNDER THAT SAME MANDATED FUNCTIONAL EQUIVALENCY, IT IS AN ADDITIONAL TASK THAT THEY HAVE TO GO THROUGH IN ASSESSING THE ACCURACY OF THEIR CAPTIONS AND THEN BEING RESPONSIBLE FOR IT. IT'S SOMETHING THAT THEY SIMPLY COULDN'T TAKE ~~OUT~~ ON. IT REQUIRES TOO MUCH WORKING MEMORY TO KIND OF HAVE THAT ONLINE ALL THE TIME.

THEN WE HAD PEOPLE ON THE OTHER SIDE OF THE SPECTRUM WHO ALSO DIDN'T USE THAT SERVICE VERY SUCCESSFULLY. ONE OF MY FAVORITES WAS I ACTUALLY CALLED ONE GENTLEMAN IN PARTICULAR TO FIND OUT HOW HIS BUTTON WORK WAS GOING AND HOW THE WHOLE THING WAS GOING FOR HIM. AND I WON'T SWEAR THE WAY HE DID, BUT HE SAID: IT DOESN'T WORK. THE DARN THING DOESN'T WORK. AND I SAID IT DOESN'T WORK? TRUST ME, IT WORKS. I KNOW IT WORKS. I BUILT IT.

AND HE SAID, NO, I'M SITTING HERE, AND I'M HITTING THIS BUTTON AND NOTHING'S HAPPENING.

AND I SAID, WELL, YOU'RE HITTING THE BUTTON WHILE YOU'RE TALKING?

AND HE SAID, YEAH.

AND I LOOKED AT THE RECORD AND HE HAD HIT IT LIKE SIX TIMES WHILE HE WAS TALKING. IT WAS SORT OF LIKE YOU WATCH SOMEONE AT AN ELEVATOR SIT THERE HAMMERING THAT BUTTON THINKING IT'S GOING TO COME ANY SECOND NOW. MAYBE THIS WILL GO FASTER. OKAY. SO WE COULD HAVE DESIGNED SOMETHING ON OR OFF. BUT THE POINT IS THESE ~~ARE~~ WERE PEOPLE TRAINED IN THEIR HOMES. WE DEMONSTRATED HOW THE SYSTEM WORKED, AND YET JUST KEEPING THAT LEVEL OF META AWARENESS, THE LEVEL OF AWARENESS ABOUT HOW THESE CAPTIONS ARE GOING IN THEIR HEAD AND THEY NEEDED TO DO SOMETHING ABOUT IT TO SWITCH FROM ONE TO ANOTHER WAS JUST TOO MUCH. SO THE ANSWER TO YOUR QUESTION IS, TECHNICALLY, IT'S POSSIBLE. PRACTICALLY, IT'S A POOR SOLUTION, I THINK.

>> DAVID BAHAR: THANKS. EARLIER I TALKED ABOUT USING MICROSOFT OFFICE TO TRANSCRIBE A PRESENTATION. NOW, THIS WAS P NOT DONE IN REALTIME. I WAS ABLE TO SLOW IT DOWN, AND IT WAS NOT 100% ACCURATE. ASR NOW IS WELL FAR AND ABOVE BEYOND THAT. BUT VISAR HOW LONG DO YOU PROJECT FOR ASR TO BECOME AS GOOD AS THE BEST HUMAN CAPTIONER?

>> VISAR BERISHA: THAT'S A VERY DIFFICULT QUESTION TO ANSWER FOR A VARIETY OF REASONS. I THINK THE MAIN ONE IS THAT BASED ON THE PRESENTATION THAT WE JUST HEARD, ACCURACY IS REALLY DIFFICULT TO QUANTIFY. IT HAS MANY, MANY DIFFERENT DIMENSIONS, RIGHT? AND SO EVEN, I HAVE SO MANY QUESTIONS AFTER HEARING THEIR PRESENTATION, I'D LOVE TO CHAT AFTERWARDS, BUT BEING ABLE TO DEFINE WHAT WE MEAN BY THE EQUIVALENT OF A HUMAN WOULD COME FIRST? I THINK WE'RE FAIRLY FAR OFF. I MEAN, THE PERFORMANCE IS IMPRESSIVE. I THINK A LOT OF THE IMPROVEMENT IN PERFORMANCE COMES ABOUT BECAUSE PEOPLE UNDERSTAND THAT THESE SYSTEMS DON'T WORK WELL, AND SO THEY ACTUALLY CHANGE, THEY MODIFY THE WAY THAT THEY SPEAK WITH WHEN THEY INTERACT WITH THEM. SO I SEE WHEN MY KIDS SPEAK WITH ALEXA, FOR EXAMPLE, THEY'LL SLOW DOWN AND SPEAK VERY DELIBERATELY, AND IT WORKS WELL. SO, OF COURSE, WE DON'T DO THAT IN CONVERSATIONAL SPEECH. AND SO ~~SO~~ IN THAT ~~P~~ SENSE, I THINK WE'RE FAR OFF. SOME OF THE ACCURACY NUMBERS THAT ARE REPORTED ARE REPORTED IN VERY SPECIFIC CASES, SO, FOR EXAMPLE, WITH READ SPEECH, YOU KNOW, IT'S KNOWN THAT WHEN HUMANS READ, IT SOUNDS VERY DIFFERENT THAN WHEN THEY CONVERSE WITH OTHERS, AND IT'S SLOWER, MORE DELIBERATE AND EASIER TO RECOGNIZE. CONVERSATIONAL SPEECH IS MUCH MORE DIFFICULT. SO I THINK WE'RE FAR AWAY. ~~BUD~~ BUT IF YOU HAD ASKED

ME, YOU KNOW, 10 YEARS AGO WHETHER THE CURRENT LEVEL OF PERFORMANCE WHERE I COULD HAVE PREDICTED THE CURRENT LEVEL OF PERFORMANCE BACK THEN, I WOULD HAVE SAID, YOU KNOW, THAT'S PROBABLY NOT FEASIBLE. SO WHAT DOES THAT MEAN? THAT MEANS THAT MAYBE 10 YEARS FROM NOW IT WILL BE. I DO KNOW THAT SOME OF THE MATHEMATICAL CONSTRUCTS THAT ARE USED TO DESIGN THESE ALGORITHMS HAVE TO IMPROVE AND BECOME A BIT MORE GENERAL BEFORE THIS IMPROVEMENT CAN COME ABOUT. SO I THINK THERE'S SOME FUNDAMENTAL IMPROVEMENTS IN TECHNOLOGY THAT HAVE TO OCCUR BEFORE WE GET, YOU KNOW, THE EQUIVALENT OF HUMAN PERFORMANCE.

>> DAVID BAHAR: THANK YOU. ASR, EVEN WHERE IT IS ~~IS~~ AMAZING, AND IT'S DIFFICULT TO PREDICT WHERE IT'S GOING, BUT ONE THING'S FOR SURE, IT'S BEING USED MORE AND MORE WIDELY EVERY DAY IN MANY DIFFERENT AREAS, AND WE SEE A LOT OF DIFFERENT PROPRIETARY SYSTEMS BEING DEVELOPED, SO ~~KRE~~ CRE AND VISAR FEEL FREE TO ANSWER, WHAT IS THE DIFFERENCE IN ASR HOW IT'S USED IN RELAY OR HOW IT'S USED IN THE GOOGLE RELAY APP OR ANY OTHER APPLICATION OF IT?

>> CRE ENGELKE: THAT'S A GREAT QUESTION. I THINK ABOUT THREE OR FOUR DIFFERENT DIRECTIONS. GOING BACK TO VISAR'S LAST QUESTION AND ANSWER, WHAT I FIND VERY INTERESTING IS THAT RIGHT NOW IN RELAY, SO THE WAY CAPTION TELEPHONE AND RELAY WORK, AND ASR WORK IS THAT YOU HAVE A HUMAN WHO ACTUALLY USES THE ASR AND THEN HAS A FULL TOOL KIT TO CORRECT AND MODIFY AND SEARCH SO THEY CAN ADD BACKGROUND NOISES, THEY CAN CORRECT WORDS THAT WEREN'T CAPTIONED CORRECTLY THE FIRST TIME. THEY CAN DO ALL SORTS OF THINGS BASED ON AS VISAR EXPLAINED EARLIER, THE WAY THEY'RE HEARING, LISTENING, THE FAMILIARITY THEY HAVE WITH THE CONTEXT OF THE SPEAKER AND WHAT NOT. SO I THINK WHAT WE'RE GOING TO ULTIMATELY FIND IS THAT IT'S NOT AS THOUGH WE HAVE HUMANS ON ONE SIDE OF THE DICHOTOMY AND MACHINES ON THE OTHER. IT'S NOT THE -- IT WAS JOHN HENRY AND THE STEEL STEAM ENGINE, IT'S NOT MAN VERSUS MACHINE IS WHAT I'M SAYING, IT'S HUMAN PLUS ASR ON THE ONE HAND VERSUS ASR WITHOUT THE HUMAN ON THE OTHER, AND THAT'S MAYBE ~~TO~~ THE ANSWER TO ONE ASPECT OF THIS QUESTION, AS THE ASR IS IMPROVED, SO, TOO WILL RELAY IN THE SENSE THAT THE HUMAN WILL ALWAYS ADD SOME ELEMENT TO THE AUGMENTED OR CORRECTED COMPONENTS OF THAT. WHEREAS THE MACHINES WILL HAVE TO GET SO GOOD THAT THEY ~~EE CLIPPINGSED~~ ECLIPSE THE COMBINED TOTAL OF HUMAN PLUS MACHINE.

THE OTHER WAY I CAN THINK OF THIS IS THAT MACHINES ARE CERTAINLY, AGAIN, TO VISAR'S EARLIER POINT, THEY'RE TRAINED ON THESE DIFFERENT MODELS, AND SO THE RELAY, MY UNDERSTANDING IS THE ~~JAS~~ VAST MAJORITY OF ASR ALGORITHMS WERE INITIALLY DEVELOPED AROUND SHORT PHRASES AND COMMANDS, BECAUSE THAT'S WHAT MOST OF THEM GET USED FOR. SO THINK ABOUT THE MILLION HOURS IN THE PAPER THAT ALEXA JUST HAD. IN FACT, I WAS AT A PRESENTATION BY THE ~~GI~~ GUY WHO INVENTED THE LABS THAT CREATE THE SPEECH MODELS FOR ALEXA A FEW MONTHS AGO AND HAD A CHANCE TO TALK TO HIM. AND HE SAID THAT THEY SPENT MILLIONS OF DOLLARS SETTING UP THESE LABS TO GET THE INITIAL GROUNDWORK FOR THEIR ALGORITHM. AND THEN WITHIN ONE WEEK OF TURNING ALEXA ON, THEY THREW AWAY ALL OF THAT DATA THAT THEY HAD SPENT MILLIONS OF DOLLARS ON, BECAUSE ~~THEZ~~ THERE'S SOME MULTIPLE OF IT ALREADY GENERATING FROM JUST RECORDING ALL THE THINGS THAT YOU SAY TO YOUR ALEXA.

BUT THE POINT IS, WHEN YOU TALK TO YOUR ALEXA, YOU TALK TO YOUR ALEXA IN A PARTICULAR WAY, SO YOU CREATE A VERY DIRECTED, VERY DEFINED SET OF UTTERANCES AND THAT CREATES ONE MODEL. BUT TELEPHONE RELAY OR CAPTION RELAY, WE DON'T HAVE THAT LUXURY OF CAPTURING ONE TYPE OF SPEECH, WE HAVE TO CAPTURE ANYTHING, ANY SPEAKER ANY TIME OF DAY, ANY ~~TOP I~~ TOPIC THINK, ANY CONDITION, ANY DEGREE OF EXCITEMENT, ANY DIALECT, ANY ACCENT, ANY DISFLUENCY, ANYTHING, SO I THINK THAT'S PART OF IT, TOO, IS WHAT IS THE DIFFERENCE? WELL, MACHINES LEARN IN THE ~~ACT~~ EXACT OPPOSITE WAY OF HUMANS. HUMANS START BY LEARNING HOW TO PROCESS SOUND AND COMMAND AND SOCIAL INTERACTIONS AND WHAT IT IS THAT, YOU KNOW, IF I CRY, THEN I GET THIS. OR THESE WORDS MEAN THAT THING. THIS IS HOW I INTERACT WITH PEOPLE. THAT'S HOW HUMANS LEARN. BUT THEN IT TAKES US UNTIL GRADUATE SCHOOL TO DEVELOP THE SKILLS TO DO ADVANCED CALCULUS AND ALL KINDS OF HIGH-LEVEL MATHEMATICS, MACHINES CAN DO THAT SORT OF PROGRAMMING IMMEDIATELY AND THAT'S WHAT THEY WERE INITIALLY DESIGNED FOR, AND IT TAKES THEM A LONG TIME TO DEVELOP THE SOCIAL MODEL THAT WE ALL GET RIGHT AWAY. THOSE ARE THE BIG DIFFERENCES BETWEEN ASRs THAT ARE USED AND THE TYPES OF RELAY THAT'S USED IN CAPTION TELEPHONE.

>> VISAR BERISHA: THERE'S A NICE RULE OF THEM. IF YOU LEARNED IT BEFORE THE AGE OF 10, A COMPUTER WILL HAVE A HARD TIME REPLICATING IT. IF YOU LEARNED IT AFTER THE AGE OF 10, IT'S EASIER. SO THAT'S WHY, YOU KNOW, HUMANS ARE MUCH BETTER AT DOING -- MACHINES ARE MUCH BETTER AT DOING CALCULUS THAN HUMANS, BUT IT'S MUCH MORE DIFFICULT FOR THEM TO UNDERSTAND SPEECH AND THE NUANCES OF SPEECH. THIS IS SOMETHING WE HAVEN'T

REALLY DISCUSSED TODAY, WHICH IS, THERE'S SO MUCH MORE TO SPEECH THAN JUST THE WORDS BEING SPOKEN. I MEAN ALL OF THE PARA LING WIS TISSUE PARTICULAR, AND THE EMOTION INHERENT IN THE ACOUSTICS, IT'S VERY DIFFICULT TO DEVELOP ALGORITHMS TO INFER THAT. SO ALTHOUGH PEOPLE ARE TRYING AND MAKING PROGRESS EVERY DAY, BUT I DON'T THINK THAT WE'RE NEAR, YOU KNOW THAT LEVEL OF PERFORMANCE YET, ANYWHERE NEAR THAT.

>> DAVID BAHAR: THANK YOU, VISAR, AND SO WE HAVE -- WE'RE ABOUT FIVE MINUTES AHEAD OF TIME NOW, AND SO I HAVE ONE LAST QUESTION FOR VISAR AND THEN WE'RE READY FOR YOUR QUESTIONS. SO IF YOUR READY. VISAR, YOU MENTIONED EARLIER THAT COMPANIES LIKE MICROSOFT, GOOGLE, APPLE, AND THE LIKE, AMAZON, THEY WORK FROM SPEECH MODELS IN ORDER TO DEVELOP THEIR ASR ALGORITHMS. DOES THAT MEAN THAT RELAY COMPANIES MIGHT START INCORPORATING ASR INTO THEIR SERVICES, AND AT WHICH POINT WHAT WOULD AK ARE I SI BE BASED ON? WHAT MODEL WOULD THEY BE USING, DO YOU THINK THAT THEY WANT FROM THESE VARIOUS COMPANIES OR SOMETHING ELSE?

>> VISAR BERISHA: THIS IS ANOTHER GOOD QUESTION. SO I THINK THIS -- YOU KNOW THE ANSWER TO THIS QUESTION, WHICH MODEL IS MOST APPROPRIATE FOR THE CPS SETTING IS A GREAT QUESTION FOR RESEARCH AND IT SEEMS LIKE IN SOME OF THE EARLIER WORK THAT YOU GUYS HAVE DONE, YOU GUYS HAVE PERHAPS LOOKED INTO THAT. IT'S CERTAINLY TRUE THAT THE SELECTION OF THE MODEL WILL IMPACT PERFORMANCE, AND THIS, AGAIN, GOES BACK TO THE QUESTIONS AROUND HOW THE MODEL WAS CALIBRATED. IF YOU LOOK AT WHAT THE COMPANIES ARE DOING, YOU CAN SORT OF INFER THE TYPE OF SPEECH SAMPLES THAT THEY HAVE AVAILABLE TO THEM. SO YOUTUBE AND GOOGLE HAVE A CLOSE PARTNERSHIP, SO A LOT OF OF THE MOLGZS THAT GOOGLE DEVELOPS ARE BASED ON AUDIO EXTRACTED FROM YOUTUBE CLIPS. IN FACT THEY'VE MADE THESE PUBLICLY AVAILABLE FOR RESEARCHERS.

AMAZON ON THE OTHER HAND, ALEXA IS A VERY POPULAR TOOL, AND THE SORT OF SPEECH SAMPLES THAT YOU GET FROM ALEXA ARE DIFFERENT FROM FOOLG, THESE ARE SHORT PHRASES, AND WITH PERHAPS MORE REVERBERATION BECAUSE PEOPLE AREN'T USING THE MICROPHONE CLOSE TO THEMSELVES TO SPEAK. SO WHICH IS SORT OF MOST APPROPRIATE FOR THE CTS SETTING? I THINK THAT'S A MATTER OF RESEARCH. IN TERMS OF ANOTHER PLAYER COMING IN AND DEVELOPING THEIR OWN ENGINE THAT'S SPECIFICALLY OPTIMIZED FOR THIS AREA, I SUPPOSE THAT'S POSSIBLE. WELL, I'LL DESCRIBE WHAT'S HAPPENING IN ACADEMIA. SO WHAT'S HAPPENING IN ACADEMIA, IS A LOT OF THE RESEARCHERS THAT FOCUSED STRICT THE

ON SPEECH RECOGNITION, IT BECOMES VERY DIFFICULT TO COMPETE WITH THE LIKES OF GOOGLE AND AMAZON, BECAUSE THE VOLUME OF DATA IS AT A SCALE THAT DOESN'T EXIST IN ACADEMIC RESEARCH LABS.

AND SO WHETHER THERE CAN BE -- I MEAN, I SUPPOSE I COULD ENVISION A THIRD-PARTY COMPANY COMING IN AND DEVELOPING THEIR OWN SPEECH RECOGNITION MODEL THAT WORKS VERY, VERY WELL FOR THIS SETTING AND THIS CONTEXT, BUT THAT'S A TALL TASK.

>> DAVID BAHAR: THANK YOU. I'D LIKE TO NOW OPEN THE FLOOR TO THE AUDIENCE FOR ANY QUESTIONS THEY HAVE. IF YOU HAVE A QUESTION, PLEASE COME FORWARD EITHER TO THE MICROPHONE OR UP TO THE FRONT NEAR THE STAGE IF YOU'RE GOING TO SIGN. IF YOU HAVE A SECOND QUESTION, I'D LIKE TO ASK YOU TO GO TO THE BACK OF THE LINE TO ASK IT SO THAT EVERYONE GETS A CHANCE TO ASK THEIR INITIAL QUESTION. RON, WOULD YOU LIKE TO START? ALSO --

>> RON : I'VE GOT TWO COMMENTS I'D LIKE TO MAKE. I'M AN EXPERT WHEN IT COMES TO ~~ICTS~~IPCTS, AND I WANT TO ENSURE THAT ALL THE DIFFERENT TEST SECTIONS THAT YOU HAVE HAD ACCURACY. I CAN SEE ACCURACY HAPPENING IN SEVERAL SITUATIONS. YOU NEVER HEAR ME COMPLAIN ABOUT ACCURACY. WHEN I HAVE HEARING AIDS ON, SO I THINK ACCURACY IS FROM A HEARING PERSON'S PERSPECTIVE. HOW DO YOU FEEL ABOUT THE LATENCY IF YOU'RE A HARD OF HEARING PERSON LIKE ME, I KNOW THERE'S A DELAY, BUT DEAF PEOPLE MAY NOT KNOW THERE IS A DELAY - DID YOU TAKE ALL THREE OF THE FACTORS INTO CONSIDERATION WHEN YOU WERE EVALUATING LATENCY AND ACCURACY?

>> CRE ENGELKE: THAT'S A TERRIFIC QUESTION AND WHAT I LIKE ABOUT IT IS IT ACTUALLY DRIVES AT WHAT YOU WERE TRYING TO GET IT FROM A USER'S PERSPECTIVE AND WHY WE DIDN'T CONSTRAIN PEOPLE'S UNDERSTANDINGS OF ACCURACY FROM THE BEGINNING OR CONSTRAIN THEIR UNDERSTANDINGS OF LATENCY. SO IF YOU REMEMBER THE GREEN AND YELLOW SLIDE THAT I HAD UP THERE, AND FORGIVE ME, I'M A BIG PACKERS FAN. (LAUGHTER) BUT THAT WAS USERS' OWN ASSESSMENTS. BUT I THINK YOU MADE ANOTHER POINT THERE THAT IS EXTREMELY IMPORTANT, WHICH IS THAT THERE'S A HUGE VARIETY OF PEOPLE WHO USE IPCTS AND CTS AND PEOPLE ARE GOING TO HAVE DIFFERENT NEEDS BASED ON HOW THEY USE THE SERVICE.

NOW, SOME OF THESE NEEDS, I THINK, ARE GOING TO BE REGULAR AND ACROSS THE WHOLE GROUP, AND SOME OF THOSE ARE GOING TO BE UNIQUE TO CERTAIN PARTS OF A GROUP, AND I THINK THIS IS WHY

WE NEED TO HAVE EXTREMELY RIGOROUS, ROBUST TESTING, SO THAT THE MAJORITY CAN BE SHOWN BUT MINORITIES CAN ALSO EMERGE. THAT IS TO SAY IF THERE'S INDIVIDUALS OR SMALL GROUPS OF PEOPLE WHO SAY, WELL, NO FOR ME I WANT IT THIS WAY OR THIS IS EXACTLY HOW I WANT MY CAPTIONING. THOSE METRICS NEED TO STAND OUT. AND THE PROBLEM WITH THE APPROACH OF THIS TESTING IS IT TENDS TO AGGREGATE ALL OF THE DIFFERENT USERS AND ALL OF THE DIFFERENT USE CASES, AND SO IT DOESN'T BRING TOGETHER -- I'M SORRY IT BRINGS TOGETHER ALL OF THESE DIFFERENT USERS INSTEAD OF WHAT ARE ALL OF THE PEAKS SDP VALLEYS IN THE RESEARCH AND HOW CAN WE MAKE A SCORECARD AS ~~DICKS YIB~~ DIXIE SAID THAT REPRESENT THE FACETS OF THE CAPTIONS THAT COME OUT SO PEOPLE CAN CHOOSE EXACTLY WHAT THEY MEAN.

>> RON: I JUST WANT TO MAKE THAT POINT CLEAR. WE HAVE 5IPCTS PROVIDERS IN THIS ROOM. WE NEED TO HAVE EVALUATION OF ASR AS IT RELATES TO RELAY. YOU MADE A COMMENT THAT AT THIS POINT IN TIME WE ARE NOWHERE NEAR BEING ABLE TO USE ASR. NOT EVEN CLOSE. THAT'S THE FIRST EVALUATION. ARE WE GETTING CLOSER TO THAT POINT? ARE WE FINDING OUT THAT THE FACT THAT WE'RE ~~WOOESING~~ WASTING OUR TIME AND SPINNING OUR WHEELS BECAUSE IT'S NOT THERE? WHERE EXACTLY ARE WE AFTER DOING THE RESEARCH? ARE WE SPINNING OUR WHEELS?

>> VISAR BERISHA: OFTENTIMES FOR THINGS AS COMPLEX AS THE ALGORITHMS THAT ARE DEVELOPED BUT ALSO AS COMPLEX AS SPEECH PRODUCTION, THERE'S MANY DEGREES OF FREEDOM THAT YOU HAVE TO TEST DURING RESEARCH. SO YOU HAVE THE VARIOUS DEFINITIONS OF ACCURACY. YOU HAVE THE VARIOUS SPEECH RECOGNITION ENGINES. YOU HAVE THE VARIOUS DEFINITIONS OF LATENCY PERCEIVED ACCURACY VERSUS PERCEIVED LATENCY. WHERE AM I GOING WITH ALL THIS? IN ORDER TO ASSESS WHAT IS THE OPTIMAL SOLUTION FOR A GIVEN CONTEXT, YOU WOULD HAVE TO EVALUATE DIFFERENT PARAMETERS OF THESE DEFINITIONS AND UNFORTUNATELY THIS TAKES A REALLY LONG TIME. SO IN TERMS OF WHERE WE ARE AFTER A SINGLE STUDY, YOU KNOW, I'LL CERTAINLY LET THEM EXPLAIN THE OUTCOMES OF THEIR STUDY BUT OFTENTIMES IN MY RESEARCH WE FIND THAT A STUDY ENDS UP SORT OF ~~REVEEG~~ REVEALING ADDITIONAL INTERESTING QUESTIONS THAT CAN BE ASKED, AND -- BUT, YOU KNOW, SIMULTANEOUSLY YOU END UP LEARNING QUITE A BIT AFTER EVERY STUDY. I THINK OVER TIME WHAT YOU END UP BUILDING IS A KNOWLEDGE BASE THAT ALLOWS YOU TO MAKE DATA-DRIVEN DECISION WHAT WHAT'S THE ~~WRITE~~ RIGHT SETTING WHEN IT SHOULD BE USED AND WHEN IT SHOULD NOT BE USED. WHO BENEFITS FROM IT AND WHAT DO WE MEAN BY BENEFIT, IT'S REALLY A VERY OPEN-ENDED QUESTION AND UNTIL THERE'S

SUFFICIENT OR CRITICAL MASSIVE RESEARCH TO BE DONE SO THAT WE CAN MAKE THESE DECISIONS, I THINK YOU HAVE TO PURSUE THESE QUESTIONS. I DON'T KNOW IF YOU GUYS HAVE MORE?

>> CRE ENGELKE: I LIKE THAT ANSWER VERY MUCH, AND I THANK YOU FOR THE QUESTION. SO I AM NOT DIRECTLY OR INTIMATELY FAMILIAR WITH ANY OF THE SYSTEMS THAT HAVE APPLIED. MY ASSUMPTION IS THAT IS AS VISAR SAID BECAUSE OF DIFFICULTIES WITH GENERATING THEIR OWN ASRs THEY ARE ~~UNWRAPPING~~ REWRAPPING SOMETHING ALREADY ON THE SHELF AND JUST PUTTING IT INTO -- YOU KNOW, PUTTING IT INTO SOME SORT OF AN APP FORM IF IT'S NOT THERE ALREADY.

AND I DON'T KNOW WHAT THEY DO, AND I CAN'T SPEAK DIRECTLY TO ANY OF THAT. BUT WHAT I DO WANT TO SAY IS COMING BACK TO SOMETHING THAT DIXIE STARTED US WITH, IS THAT THE GOAL HERE IS REALLY DEVELOPING DATA-DRIVEN DECISION-MAKING AND THE TYPES OF QUALITATIVE AND QUANTITATIVE METRICS WOULD ALLOW EVERYONE TO FEEL CONFIDENT THAT WHAT THEY'RE USING IS GOING TO BE THE HIGHEST POSSIBLE QUALITY.

SO I DON'T HAVE ANY DIRECT ACCESS TO ANY OF THE USE OF THIS INDEPENDENT ASR AND IPCTS OTHER THAN RESEARCH THAT WE'VE DONE, BUT THE GOAL I THINK SHOULD ALWAYS BE MAKING DECISIONS ON THE DATA AND LETTING THAT TAKE REGULATION OR WHATEVER ELSE, APPROVALS, FOR.

>> DIXIE ZIEGLER: QUALITY METRICS, RIGHT? THAT'S HOW WE KNOW. AND WE NEED TO BE ABLE TO HAVE SOME WAY TO BE ABLE TO MEASURE IT, AND ALL OF THAT IS TRUE. WE NEED THE RIGHT FRAMEWORK, SAMPLING, TESTING, THE RIGHT SCORING METRICS, AND WE NEED TO BE ABLE TO DO THAT FOR EVERYBODY AND UNDERSTAND WHAT THOSE -- WHAT THAT IS TO ENSURE THAT WE HAVE THE RIGHT QUALITY PROTECTIONS IN PLACE TO ENSURE THAT CONSUMERS ARE GETTING THE BEST THAT THEY CAN GET.

>> DAVID BAHAR: THANK YOU, RON, FOR YOUR QUESTION. AND NOW WOULD YOU PLEASE INTRODUCE YOURSELF BEFORE ASKING YOUR QUESTION.

>> I'M LORI FROM WYOMING, AND I HAVE TWO QUESTIONS, IF IT'S OKAY. IF NOT, THEN STOP ME. MY FIRST QUESTION RELATES TO THE STUDY WHERE YOU LOOKED AT USER EXPERIENCE EMPOWERMENT AND YOU COMPARED IT WITH ACCURACY AND LATENCY. I'M CURIOUS IF YOU COLLECTED AND LOOKED AT CORRELATIONS WITH DEMOGRAPHICS OF THE USERS, IF YOU BROKE THAT DOWN ANY BY AGE OR OTHER

CHARACTERISTICS? WE, WE ALSO DO THE EQUIPMENT PROGRAM, AND WE FIND THAT THERE IS USER CHARACTERISTICS THAT MAKE INDIVIDUALS MORE LIKELY TO ADOPT AND UTILIZE THE TECHNOLOGY THAN OTHERS. FOR EXAMPLE, AGE AND POSSIBLY SOME COGNITIVE DECLINES OFTENTIMES, EVEN THOUGH THEIR HEARING LEVEL WOULD INDICATE A GOOD ~~ICPTS~~ IPCTS USER, IT DOESN'T WORK OUT BECAUSE OF STRUGGLES LIKE THAT, SO DID YOU BREAK DOWN ANY DATA BY DEMOGRAPHICS OF THE USER?

>> CRE ENGELKE: WE DID, ACTUALLY, AND IN FACT, ALL OF THE SLIDES THAT I WENT THROUGH TODAY, AND MORE, WHICH INCLUDE DEMOGRAPHIC SLIDES WERE ALL PUBLISHED AS PART OF AN EX PARTE.

>> DIXIE ZIEGLER: BACKS IN DECEMBER OF 2018, SO IT'S PRETTY EASY TO FIND ON THE FCC'S WEBSITE, SO WE DID LIST OUT THE DIFFERENT DEMOGRAPHICS, AND SO JUST A QUICK GENERAL ANSWER HERE, THESE WERE USERS WHO HAD BEEN USING THE SERVICE FOR A WHILE AND HAD GOOD EXPERIENCE ON IT, SO THAT WAS ONE OF OUR QUALIFYING INDICATORS IS TO MAKE SURE THAT THEY HAD BEEN A USER FOR AT LEAST SEVERAL MONTHS, IF NOT LONGER AND WERE USING THE SERVICE WITH REGULARITY. AND THEN THERE WERE A VARIETY OF AGES AND THE LIKE, BUT, YES, THERE'S ALTS BIT MORE DETAIL IN THE EX PARTE WE FILED BACK IN DECEMBER OF 2018.

>> CRE ENGELKE: AND WE WERE TRYING TO AVOID ANY ADDITIONAL SORT OF EX PARTE. (LAUGHTER) SO WE ~~WEE CYCLED~~ RECYCLED OUR SLIDES.

>> MY SECOND QUESTION IS YOU SAID YOU ALLOWED INDIVIDUALS TO SWITCH FROM TRADITIONAL ~~ICPTS~~ IPCTS TO ASR AS A USER CHOICE, DID YOU HAVE ANY OTHER CRITERIA WHERE IT THEY'D MADE THAT SWITCH, FOR EXAMPLE CERTAIN TERMINOLOGY OR CERTAIN SITUATIONS THAT WOULD SAY, NOW LET'S MOVE FROM ASR INTO IPCTS OR WAS IT ALWAYS USER CHOICE.

>> CRE ENGELKE: SO IN THAT PARTICULAR STUDY THAT I TALKED ABOUT IT WAS ALL USER CHOICE. AND I APPRECIATE YOUR QUESTION. BUT IN THAT PARTICULAR STUDY WE ONLY WERE LOOKING AT THE USER'S CAPACITY TO RECOGNIZE THE QUALITY OF CAPTIONS IN THE MIDST, SORRY, RECOGNIZE AND TAKE ACTION IN THE MIDST OF A CAPTIONING TELEPHONE CALL, AND SO WE WERE REALLY FOCUSED ON THAT, AND WE DID IT IN THE WAY WE DID BECAUSE WE DIDN'T WANT TO CREATE ADDITIONAL OUTSIDE OR INCLUDE ADDITIONAL OUTSIDE BIASES, SUCH AS, FOR EXAMPLE, SOME PEOPLE MAY FEEL STRONGLY THAT THEY WANT ASR BECAUSE THEY WANT SOME L THAT THE ASR

GIVES THEM. SOME PEOPLE DON'T WANT ASR BECAUSE THEY ARE CONCERNED ABOUT THE ASR ~~BINGZ~~ ~~IN~~ PRIVACY. WE DID IN A BLIND WAY, AND WE FOCUSED ONLY ON THE CAPTIONS THE ACCURACY AND THE DELAY OF THE CAPTIONS AND NOT ON ANY OUTSIDE PARAMETERS, BECAUSE WE WERE SIMPLY LOOKING AT THEIR ABILITY TO ASSESS AND TAKE ACTION IN THAT REGARD.

>> DIXIE ZIEGLER: IT DOESN'T MAKE YOUR QUESTION WRONG, BUT THAT WAS THE QUESTION WE ANSWERED ON.

>> CRE ENGELKE: UNFORTUNATELY TO GET THE LEVELS OF ~~ASSURE AT~~ ASSURITY THIS FOR THE TYPES OF RESEARCH WE WERE DOING, YOU HAVE TO ASK SMALL-BOUNDED QUESTIONS AND HOPEFULLY THEY BUILD UP.

>> DIXIE ZIEGLER: THAT BRINGS UP THE POINT OF HOW ONE QUESTION LEADS TO 10 MORE.

>> DAVID BAHAR: OKAY. WE HAVE 8 MINUTES REMAINING, IF ANYBODY WOULD LIKE TO ASK ANOTHER QUESTION. GOING ONCE, GOING TWICE, GOING, GOING, GONE. OKAY.